PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 39/395, C07H 21/02, 21/04, C07K 14/435, 16/30, C12Q 1/68, G01N 33/574 — A1

(11) International Publication Number: WO 98/22139

(43) International Publication Date: 28 May 1998 (28.05.98)

(21) International Application Number: PCT/US97/21807

(22) International Filing Date: 19 November 1997 (19.11.97)

(30) Priority Data:
60/031,711 22 November 1996 (22.11.96) US

(71) Applicant: LA JOLLA INSTITUTE FOR ALLERGY AND IMMUNOLOGY [US/US]; 10355 Science Center Drive, San Diego, CA 92121 (US).

(72) Inventors: HUFLEJT, Margaret, E.; 8536 Via Mallorca Drive #A, La Jolla, CA 92037 (US). LIU, Fu-Tong; 4351 Mensha Place, San Diego, CA 92130 (US).

(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DETECTION OF GALECTIN-4 IN HUMAN TUMORS

(57) Abstract

This invention is based on the discovery of the human galectin-4 gene, and the discovery that human galectin-4 is expressed in breast cancer. The invention provides an isolated polynucleotide encoding human galectin-4, and the human galectin-4 polypeptide. Antibodies which bind human galectin-4 polypeptide and formulations for administration of these antibodies are also disclosed. The present invention provides a method of diagnosing or determining the prognosis of cell proliferative disorder, such as breast cancer, associated with the human galectin-4. A method of determining the presence of metastases in a sample from a subject by determining the presence or absence of galectin-4 is also disclosed. A method of treating a subject having, or at risk of having, a human galectin-4 associated disorder, such as breast cancer, is also provided. A kit useful for detecting the presence of human galectin-4 polypeptide or polynucleotide in a sample from a subject having a human galectin-4-associated disorder is disclosed. A method of identifying compounds which affect human galectin-4 expression is also provided. Transgenic nonhuman animals having a transgene encoding human galectin-4 are also described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

DETECTION OF GALECTIN-4 IN HUMAN TUMORS

Field of the Invention

The present invention relates to generally to antigens which are differentially expressed in human breast tumor cells as compared to normal cells, and to the diagnosis, study, prevention and treatment of disease based on this differential expression.

Background of the Invention

Breast cancer is the most common form of malignant disease in women in western countries; in the United States it is the most common cause of death of women between the ages of 40 and 55. This disease will develop in about twelve percent of woman in the United States; about seventy percent of the cases can be cured (Forrest, A.P., 1990, "Screening and breast cancer incidence," *J. Natl. Cancer Inst.* 82:1525).

The characteristics most associated with increase relative risk of breast cancer include: (1) a first-degree female relative who has had breast cancer, (2) prior breast cancer, (3) nulliparity, (4) age greater than 30 years at first pregnancy, (5) early onset of menstruation or late menopause, and (6) radiation exposure (Kelsey, J.L., 1979, "A review of the epidemiology of human breast cancer," *Epidemiol. Rev.* 1:47). Prognosis is an assessment of the risk of relapse of breast cancer on the basis of anatomic, histologic, cytologic and biologic factors. If there is no evidence of systemic spread, the number of involved axillary lymph nodes, the histological grade of the tumor, and to a lesser extent, the size of the primary tumor and the hormone receptor status are the most important indicators of prognosis (Sigurdsson, H., *et al.*, 1990, "Indicators of prognosis in node-negative breast cancer, " *New Engl. J. Med.* 322:1045).

Genetic analyses of families with a high incidence of breast cancer suggest that specific mutations within the genome can be responsible for breast cancer. The

nucleotide and amino acid sequences associated with breast tumors can serve as genetic markers of inheritable breast cancer. Genetic changes on chromosome 17 are the most frequently identified events associated with breast tumors. At least four markers on chromosome 17 have been identified: p53 on 17p13.1, regions of loss of heterozygosity (LOH) on 17p13.3 and 17q12-qter, the breast/ovarian cancer locus (BRCA-1) on 17q21, and a fourth breast cancer growth suppressor gene on chromosome 17 (Casey *et al.*, 1993, *Hum. Molec. Genet.* 2:1921-1927).

Such genetic markers can also be useful in identifying patients susceptible to breast cancer. For example, the genetic marker BRCA-1 has been linked to a susceptibility of developing breast and/or ovarian cancer at a young age in a number of families (Hall *et al.*, 1990, *Science* 250:1684-1689; Solomon *et al.*, 1991, *Cytogenet. Cell. Genet.* 58:686-738). The cumulative risks of developing breast cancer associated with the BRCA-1 marker are 50% at 50 years and 82% at 70 years (Easton *et al.*, 1993, *Am. J. Hum. Genet.* 52:678-701). However, since the gene encoding BRCA-1 has not been cloned or sequenced, identification of an individual carrier of BRCA-1 is not possible without use of linkage analysis. Linkage analysis is generally not feasible in clinical practice since the genetic epidemiology required is tedious, if not impossible, in most cases (Kent *et al.*, 1995, *Europ. J. Surg. Oncol.* 21:240-241).

It has been proposed that the development of breast cancer is associated with multiple genetic changes associated with alterations in expression of specific genes. Several studies have indicated that the expression markers erbB-2, cathepsin D, cyclin D, p53, Rb, and c-myc may have prognostic benefit, although the evidence has been obtained on a limited number of patients in a limited number of studies. In addition, two markers, nm23, and angiogenesis, have been shown in small cohorts to have predictive value (Porter-Jordan *et al.*, 1994, *Hematol. Oncol. Clin. North Am.* 8:73-100).

The specific set of alterations within a given cancer may provide information on susceptibility to treatment and prognosis. Thus, the discovery of nucleotide sequences and polypeptides encoding proteins associated with breast

cancer would satisfy a need in the art by providing new means of diagnosing and treating breast cancer.

Summary of the Invention

This invention is based on the discovery of the human galectin-4 gene, and the discovery that human galectin-4 is expressed in breast cancer. In a first embodiment, this invention provides an isolated polynucleotide encoding human galectin-4. The invention features (1) human galectin-4 as encoded by SEQ ID NO:1, (2) SEQ ID NO:1 where T can also be a U, (3) nucleic acid sequences complementary to SEQ ID NO:1, and (4) fragments of (1)-(3) that are at least 15 bases in length and hybridize to DNA which encodes human galectin-4. In a second embodiment, the invention provides isolated human galectin-4 polypeptide. Antibodies which bind human galectin-4 polypeptide and formulations for administration of these antibodies are also disclosed.

In a third embodiment, the invention provides a method for identifying a subject having a disorder associated with human galectin-4. In another embodiment, a method of treating a subject having or at risk of having a human galectin-4 associated disorder is provided. In yet another embodiment, a method for determining the prognosis of a subject diagnosed with cancer by determining the presence or absence of human galectin-4 is provided. A method of determining the presence of metastases in a sample from a subject by determining the presence or absence of galectin-4 is also disclosed.

In a further embodiment, transgenic nonhuman animals having a transgene encoding human galectin-4 are also described.

In an additional embodiment, a kit useful for detecting the presence of human galectin-4 polypeptide in a sample from a subject having a human galectin-4-associated disorder is provided. A kit useful for the detection of polynucleotide encoding human galectin-4 in a subject having a human galectin-4 associated disorder is also provided.

In yet another embodiment, a method of identifying compounds which affect human galectin-4 expression is also disclosed.

Brief Description of the Drawings

Fig. 1. shows the expression of both galectin-3 and galectin-4 by T84 cells. Lane a: galectin fraction visualized by Coomassie staining; lane b: autoradiogram of the gel shown in 1a after 16 h of exposure; lane c: immunoblot with rabbit antiserum to the C-terminal domain of rat intestinal galectin-4; lane d: autoradiogram of an anti-galectin-3 immunoprecipitate after 16 h of exposure.

Fig. 2. is the nucleotide and deduced amino acid sequence of human galectin-4. Fig. 2A: cDNA sequence with deduced amino acid sequence. Fig. 2B: Comparison of deduced amino acid sequences of human galectin-4 from T84 cells (hGal-4), rat galectin-4 (rGal-4; Oda *et al.*, 1993), and porcine galectin-4 (pGal-4; Chiu *et al.*, 1994). Dots signify the same residue as in hGal-4, dashes represent gaps introduced for alignment. The residues that are conserved in most galectins, and known to interact with bound lactose (Barondes *et al.*, 1994) are indicated by asterisks.

Fig. 3 shows the localization of galectin-3 and galectin-4 in a confluent monolayer of T84 cells by laser scanning confocal microscopy (XZ sections) after double immunolabeling. Panel a shows cells fixed with PLP, while panel b shows cells fixed with paraformaldehyde. Apical galectin-3 is visible as green fluorescence (fluorescein-conjugated streptavidin) at the top of each panel, and the basal galectin-4 is marked by the presence of a red immunofluorescence (Texas Red-conjugated secondary antibody) at the bottom. All bars, 25 μm.

Fig. 4 shows the localization of galectin-3 and galectin-4 in a confluent monolayer of T84 cells by laser scanning confocal microscopy (XY sections). The figure shows XY sections at different levels of the same sample as shown in Fig. 3a. a: apical surface, 26 μm above the glass level; b: optical section in the middle of the monolayer at 13 μm above glass level; c: bottom of the monolayer, 1 μm above the glass level; d: higher magnification of the apical surface from panel a, showing

saponin-resistant accumulations of galectin-4 and galectin-3, with the characteristic morphology. All bars, 25 μm.

Fig. 5 show the localization of galectin-4 and cytokeratin in calcium-depleted confluent T84 monolayers (XZ sections). Confluent T84 cell monolayers were treated with EGTA for 15 min. (Fig. 5a), or for 30 min. (Fig. 5b), and then fixed but not permeabilized prior to incubation with antibodies. Galectin-4 (red, Texas Red-conjugated secondary antibody) was visualized simultaneously with cytokeratin (green, fluorescein-conjugated streptavidin). Bar, 25 μm.

Fig. 6 shows the localization of galectin-4 and E-cadherin in calcium-depleted confluent T84 monolayers (XY sections). Galectin-4 (red, Texas Red-conjugated secondary antibody) was visualized simultaneously with the intracellular domain of E-cadherin (green, fluorescein-conjugated streptavidin). Fig. 6a: optical XY-section at the apical surface, 24 μm above the substrate level; Fig. 6b: 16 μm above the substrate level; Fig. 6c: 8 μm above the substrate level; Fig. 6d: 1 μm above the substrate level. Bar, 25 μm.

Fig. 7 shows the localization of galectin-3 (green fluorescein-conjugated streptavidin) and galectin-4 (red, Texas Red-conjugated antibody) in subconfluent T84 cell cultures at different times after seeding. Fig. 7 a: XZ-section of attaching cells, 2 hours after seeding, showing accumulation of galectin-4 (arrows) at the cell-substrate contact sites. Fig. 7b: XY-section collected in the middle of the cell on the top of the cluster shown in panel a showing segregation of the two galectins. Fig. 7c: XY section of dividing cells 12 hours after seeding, 2 μm above the substrate. Fig. 7d: XY section of cell colony 24 hours after seeding, 2 μm above the substrate, small arrow points to the galectin-4-rich domain at the tip of a lamellipodium, and big arrows point to the membrane fragments often seen adhering to the uncoated glass. Fig. 7c: XY section of lamellipodia at the edge of a serum-starved cell colony 48 hours after seeding, 2 μm above the substrate. Fig. 7f: XY section 2 μm above the substrate, and XZ-section (inset) of lamellipodium at the edge of a cell colony 48 hours after seeding, 15 min. after addition of 5% fetal calf serum, N-cell nucleus. All bars, 25 μm.

Fig. 8 show the localization of galectin-3 (Fig. 8a, 8c, and 8e) and galectin-4 (Fig. 8b, 8d, and 8f) in lamellipodia of T84 cell colonies 48 hours after seeding. Notice the highest concentration of galectins in the lamellipodia of the cells at the colony edges. Bars in Fig. 8a and 8b are 100 μm, and bars in Figs. 8c-8f are 25 μm.

Fig. 9 show the adhesion of T84 cells to surface-immobilized rat recombinant galectin-4. The results are presented as the mean percent adhering cells in five to eight replicates +/- SD.

Fig. 10 show the immunohistochemical staining patterns for galectin-4 in normal and neoplastic breast tissues. Fig. 10A: Normal tissue adjacent to the tumor area with no reactivity for galectin-4 in either the ductal epithelium and stroma (original magnification=200X); Fig. 10B: Normal tissue on the left side (N) shows no reactivity for galectin-4, while the adjacent tumor area (T), shows reactivity (original magnification=100X); Fig. 10C: In situ carcinoma cells and surrounding invasive cancer show galectin-4 staining, while stroma does not show galectin-4 staining (original magnification=200X); Fig. 10D: Cytoplasm of in situ carcinoma cells show reactivity for galectin-4, while the nuclei are nonreactive (original magnification=400X); Fig. 10E: In situ cribriform carcinoma and admixed invasive carcinoma show positive staining for galectin-4 (original magnification=100X); Fig. 10F: Nonreactive stroma surrounding strongly reactive in situ cribriform tumor (original magnification=400X); Fig. 10G: Aggregates of small lymphocytes in the tumor tissue are negative for galectin-4, whereas two infiltrating invasive tumor cells (arrows) are reactive (original magnification=400X); Fig. 10H: Single invasive tumor cell(arrow) reactive for galectin-4. Note surrounding fat tissue is nonreactive (original magnification=400X)

Description of the Preferred Embodiments

The present invention is based on the discovery of human galectin-4. This invention provides isolated polypeptides and polynucleotides encoding human galectin-4 polynucleotides. Antibodies which bind to galectin-4 are also provided. In

one embodiment, the invention provides a method for detection of a disorder associated with galectin-4. In another embodiment, the invention provides a method for treating a disorder associated with galectin-4.

Rat galectin-4, a 36 kD protein, is a novel type of animal lectin containing two tandem carbohydrate-binding domains in a single polypeptide (Oda *et al.*, 1993), thus being capable of bridging two ligands. Carbohydrate-binding specificity of both carbohydrate-binding domains of rat galectin-4 has been investigated (Oda *et al.*, 1993). It was determined that both domains of galectin-4 recognize the same major carbohydrate structural determinants of β-galactosides as do other members of the galectin family. However, it was also shown that the distinct differences in the affinity for more complex carbohydrate structures between the two domains are likely predisposing galectin-4 to simultaneously bind two ligands containing different complex carbohydrate chains. Galectin-4 is found in the alimentary tract (Leffler *et al.*, 1989; Chiu *et al.*, 1992; Chiu *et al.* 1994; Wasano and Hirakawa, 1995; Tardy *et al.*, 1995)

The present invention relates to human galectin-4 and to the use of human galectin-4 antibodies, nucleic acid sequences, and amino acid sequences in the study, diagnosis, prevention and treatment of disease, including tumorigenesis. The extent and importance of changes in glycosylation patterns of both cell surface and extracellular matrix proteins in the tumor tissues is gaining increased attention (Rhodes, 1996; Taylor-Papadimitriou and Epenetos, 1994; Kim, 1992). These glycosylation changes are apparently an integral component of dynamics of the cell-ECM interactions, and are part of the pathobiological "reprogramming" of the neoplastic cell. Most of ECM proteins produced by tumorigenic cell lines or isolated from neoplasms carry additional glycosylation, and/or are glycosylated *de novo* (Mandel *et al.*, 1992).

Human galectin-4 may have application for the treatment of various types of cancer. The galectins have been reported to be involved in many disorders. An example of a galectin-3 associated disorder is expression in human T cell lines in response to infection with human T-lymphotropic virus HTLV-I (Hsu *et al.*, 1996).

Quantitative changes in galectin-3 expression have been observed in colorectal carcinomas, where enhanced galectin-3 expression correlated with decreased long term patient survival (Schoeppner *et al.*, 1995). However, progression of human breast cancer was associated with decreased expression of galectin-3 (Castronovo *et al.*, 1996). Two reports show also tumor-related changes in expression of two galectin-4-like 36 kD proteins. One of these reports shows that galectin-4-like protein is induced in prostate tumors, and that this protein is shed from prostate cancer cells (Su *et al.*, 1996). In the other study, galectin-4-like protein (but different from the prostate tumor galectin) has been identified as the major tumor antigen in patients with Hodgkin disease, and role of this protein in the cell-cell adhesive interactions between neoplastic Hodgkin and Reed-Sternberg cells and their surrounding T-cells has been proposed.

The invention provides substantially purified human galectin-4 polypeptide. Preferable, human galectin-4 has an amino acid sequence set forth in SEQ ID NO:2. The term "substantially purified" as used herein refers to a polypeptide which is substantially free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated. One skilled in the art can purify human galectin-4 using standard techniques for protein purification. The substantially pure polypeptide will yield a single major band on a non-reducing polyacrylamide gel. The purity of the human galectin-4 polypeptide can also be determined by amino-terminal amino acid sequence analysis. Human galectin-4 polypeptide includes functional fragments of the polypeptide, as long as the activity of human galectin-4 remains. Smaller peptides containing the biological activity of human galectin-4 are included in the invention. Included in the invention are polypeptides having an amino acid sequence which is at least 85% identical to SEQ ID NO:2, more preferably 90% identical to SEQ ID NO:2.

The invention provides polynucleotides encoding the human galectin-4 protein. These polynucleotides include DNA, cDNA and RNA sequences which encode human galectin-4. It is understood that all polynucleotides encoding all or a portion of human galectin-4 are also included herein, as long as they encode a

polypeptide with human galectin-4 activity. Such polynucleotides include naturally occurring, synthetic, and intentionally manipulated polynucleotides. For example, human galectin-4 polynucleotide may be subjected to site-directed mutagenesis. The polynucleotide sequence for human galectin-4 also includes antisense sequences. The polynucleotides of the invention include sequences that are degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included in the invention as long as the amino acid sequence of Human galectin-4 polypeptide encoded by the nucleotide sequence is functionally unchanged.

Specifically disclosed herein is a polynucleotide sequence containing the human galectin-4 gene. Preferably, the human galectin-4 nucleotide sequence is SEQ ID NO:1. The term "polynucleotide" or "nucleic acid sequence" refers to a polymeric form a nucleotides at least 10 bases in length. By "isolated polynucleotide" is meant a polynucleotide that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g. a cDNA) independent of other sequences. The nucleotides of the invention can be ribonucleotides, deoxyribonucleotides, or modified forms of either nucleotide. The term includes single and double forms of DNA.

The polynucleotide encoding human galectin-4 includes Figure 1 (SEQ ID NO:1), as well as nucleic acid sequences complementary to SEQ ID NO:1. A complementary sequence may include an antisense nucleotide. When the sequence is RNA, the deoxynucleotides A, G, C, and T of SEQ ID NO:1 are replaced by ribonucleotides A, G, C, and U, respectively. Also included in the invention are fragments of the above-described nucleic acid sequences that are at least 15 bases in length, which is sufficient to permit the fragment to selectively hybridize to DNA that encodes the protein of SEQ ID NO:2 under physiological conditions or a close family

member of human galectin-4. The term "selectively hybridize" refers to hybridization under moderately or highly stringent conditions which excludes non-related nucleotide sequences.

In nucleic acid hybridization reactions, the conditions used to achieve a particular level of stringency will vary, depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (*e.g.*, GC v. AT content), and nucleic acid type (*e.g.*, RNA v. DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is immobilized, for example, on a filter.

An example of progressively higher stringency conditions is as follows: 2 x SSC/0.1% SDS at about room temperature (hybridization conditions); 0.2 x SSC/0.1% SDS at about room temperature (low stringency conditions); 0.2 x SSC/0.1% SDS at about 42°C (moderate stringency conditions); and 0.1 x SSC at about 68°C (high stringency conditions). Washing can be carried out using only one of these conditions, *e.g.*, high stringency conditions, or each of the conditions can be used, *e.g.*, for 10-15 minutes each, in the order listed above, repeating any or all of the steps listed. However, as mentioned above, optimal conditions will vary, depending on the particular hybridization reaction involved, and can be determined empirically.

Minor modifications of the recombinant human galectin-4 primary amino acid sequence may result in proteins which have substantially equivalent activity as compared to the human galectin-4 polypeptide described herein. Such modifications may be deliberate, as by site-directed mutagenesis, or may be spontaneous. All of the polypeptides produced by these modifications are included herein as long as the biological activity of human galectin-4 still exists. Further, deletion of one or more amino acids can also result in a modification of the structure of the resultant molecule without significantly altering its biological activity. This can lead to the development of a smaller active molecule which would have broader utility. For example, one can remove amino or carboxy terminal amino acids which are not required for human galectin-4 biological activity.

The nucleotide sequence encoding the human galectin-4 polypeptide of the invention includes the disclosed sequence and conservative variations thereof. The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as thc substitution of arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine, and the like. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid provided that antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide.

DNA sequences encoding human galectin-4 can be expressed *in vitro* by DNA transfer into a suitable host cell. "Host cells" are cells in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However, such progeny are included when the term "host cell" is used. Methods of stable transfer, meaning that the foreign DNA is continuously maintained in the host, are known in the art.

In the present invention, the human galectin-4 polynucleotide sequences may be inserted into an expression vector. The term "expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of the human galectin-4 genetic sequences. Polynucleotide sequence which encode human galectin-4 can be operatively linked to expression control sequences. “Operatively linked” refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. An expression control sequence operatively linked to a coding sequence is ligated such that expression of the coding sequence is achieved under conditions compatible with the expression control sequences. As used herein, the term “expression control sequences” refers to nucleic acid sequences that regulate the

expression of a nucleic acid sequence to which it is operatively linked. Expression control sequences are operatively linked to a nucleic acid sequence went the expression control sequences control and regulate the transcription and, as appropriate, translation of the nucleic acid sequence. thus expression control sequences can include appropriate promoters, enhancers, transcription terminators, as start codon (i.e., ATG) in front of a protein-encoding gene, splicing signal for introns, maintenance of the correct reading frame of that gene to permit proper translation of mRNA, and stop codons. The term "control sequences" is intended to included, at a minimum, components whose presence can influence expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. Expression control sequences can include a promoter.

By "promoter" is meant minimal sequence sufficient to direct transcription. Also included in the invention are those promoter elements which are sufficient to render promoter-dependent gene expression controllable for cell-type specific, tissue-specific, or inducible by external signals or agents; such elements may be located in the 5' or 3' regions of the gene. Both constitutive and inducible promoters, are included in the invention (see *e.g.*, Bitter *et al.*, 1987, *Methods in Enzymology* 153:516-544). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage γ, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like may be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus 7.5K promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the nucleic acid sequences of the invention.

In the present invention, the polynucleotide encoding human galectin-4 may be inserted into an expression vector which contains a promoter sequence which facilitates the efficient transcription of the inserted genetic sequence of the host. The expression vector typically contains an origin of replication, a promoter, as well as

specific genes which allow phenotypic selection of the transformed cells. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in bacteria (Rosenberg *et al.*, 1987, *Gene*, 56:125), the pMSXND expression vector for expression in mammalian cells (Lee and Nathans, 1988, *J. Biol. Chem.*, 263:3521) and baculovirus-derived vectors for expression in insect cells. The DNA segment can be present in the vector operably linked to regulatory elements, for example, a promoter (e.g., T7, metallothionein I, or polyhedrin promoters).

Polynucleotide sequences encoding human galectin-4 can be expressed in either prokaryotes or eukaryotes. Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Biologically functional viral and plasmid DNA vectors capable of expression and replication in a host are known in the art. Such vectors are used to incorporate DNA sequences of the invention.

By "transformation" is meant a permanent genetic change induce in a cell following incorporation of new DNA (i.e. DNA exogenous to the cell). Where the cell is a mammalian cell, the permanent genetic change is generally achieved by introduction of the DNA into the genome of the cell.

By "transformed cell" is meant a cell into which (or into an ancestor of which has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding human galectin-4. Transformation of a host cell with recombinant DNA may be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *E. coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method using procedures well known in the art. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell if desired.

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as

microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used. Eukaryotic cells can also be cotransformed with DNA sequences encoding the human galectin-4 of the invention, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein (see for example, *Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982).

Isolation and purification of microbial expressed polypeptide, or fragments thereof, provided by the invention, may be carried out by conventional means including preparative chromatography and immunological separations involving monoclonal or polyclonal antibodies.

ANTIBODIES

The human galectin-4 polypeptides of the invention can also be used to produce antibodies which are immunoreactive or bind to epitopes of the human galectin-4 polypeptides. Antibody which consists essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations are provided.

The preparation of polyclonal antibodies is well-known to those skilled in the art. See, for example, Green *et al.*, *Production of Polyclonal Antisera*, in: Immunochemical Protocols (Manson, ed.), pages 1-5 (Humana Press 1992); Coligan *et al.*, *Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters*, in: Current Protocols in Immunology, section 2.4.1 (1992), which are hereby incorporated by reference.

The preparation of monoclonal antibodies likewise is conventional. See, for example, Kohler & Milstein, 1975, *Nature* 256:495; Coligan *et al.*, sections 2.5.1-2.6.7; and Harlow *et al.*, in: Antibodies: a Laboratory Manual, page 726 (Cold Spring Harbor Pub. 1988), which are hereby incorporated by reference. Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising an

antigen, verifying the presence of antibody production by removing a serum sample, removing the spleen to obtain B lymphocytes, fusing the B lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures. Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography. *See, e.g.*, Coligan *et al.*, sections 2.7.1-2.7.12 and sections 2.9.1-2.9.3; Barnes *et al.*, *Purification of Immunoglobulin G (IgG)*, in: Methods in Molecular Biology, Vol. 10, pages 79-104 (Humana Press 1992).

Methods of *in vitro* and *in vivo* multiplication of monoclonal antibodies are well known to those skilled in the art. Multiplication *in vitro* may be carried out in suitable culture media such as Dulbecco's Modified Eagle Medium or RPMI 1640 medium, optionally replenished by a mammalian serum such as fetal calf serum or trace elements and growth-sustaining supplements such as normal mouse peritoneal exudate cells, spleen cells, bone marrow macrophages. Production *in vitro* provides relatively pure antibody preparations and allows scale-up to yield large amounts of the desired antibodies. Large scale hybridoma cultivation can be carried out by homogenous suspension culture in an airlift reactor, in a continuous stirrer reactor, or in immobilized or entrapped cell culture. Multiplication *in vivo* may be carried out by injecting cell clones into mammals histocompatible with the parent cells, *e.g.*, osyngeneic mice, to cause growth of antibody-producing tumors. Optionally, the animals are primed with a hydrocarbon, especially oils such as pristane (tetramethylpentadecane) prior to injection. After one to three weeks, the desired monoclonal antibody is recovered from the body fluid of the animal.

Therapeutic applications for antibodies disclosed herein are also part of the present invention. For example, antibodies of the present invention may also be derived from subhuman primate antibody. General techniques for raising therapeutically useful antibodies in baboons can be found, for example, in

Goldenberg *et al.*, International Patent Publication WO 91/11465 (1991) and Losman *et al.*, 1990, *Int. J. Cancer* 46:310, which are hereby incorporated by reference.

Alternatively, a therapeutically useful anti-human galectin-4 antibody may be derived from a "humanized" monoclonal antibody. Humanized monoclonal antibodies are produced by transferring mouse complementarity determining regions from heavy and light variable chains of the mouse immunoglobulin into a human variable domain, and then substituting human residues in the framework regions of the murine counterparts. The use of antibody components derived from humanized monoclonal antibodies obviates potential problems associated with the immunogenicity of murine constant regions. General techniques for cloning murine immunoglobulin variable domains are described, for example, by Orlandi *et al.*, 1989, *Proc. Nat'l Acad. Sci. USA* 86:3833, which is hereby incorporated in its entirety by reference. Techniques for producing humanized monoclonal antibodies are described, for example, by Jones *et al.*, 1986, *Nature* 321:522; Riechmann *et al.*, 1988, *Nature* 332:323; Verhoeyen *et al.*, 1988, *Science* 239:1534; Carter *et al.*, 1992, *Proc. Nat'l Acad. Sci. USA* 89:4285; Sandhu, 1992, *Crit. Rev. Biotech.* 12:437; and Singer *et al.*, 1993, *J. Immunol.* 150:2844, which are hereby incorporated by reference.

Antibodies of the invention also may be derived from human antibody fragments isolated from a combinatorial immunoglobulin library. See, for example, Barbas *et al.*, 1991, in: Methods: a Companion to Methods in Enzymology, Vol. 2, page 119; Winter *et al.*, 1994, *Ann. Rev. Immunol.* 12:433, which are hereby incorporated by reference. Cloning and expression vectors that are useful for producing a human immunoglobulin phage library can be obtained, for example, from STRATAGENE Cloning Systems (La Jolla, CA).

In addition, antibodies of the present invention may be derived from a human monoclonal antibody. Such antibodies are obtained from transgenic mice that have been "engineered" to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain loci are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy and light chain loci. The transgenic

mice can synthesize human antibodies specific for human antigens, and the mice can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described by Green *et al.*, 1994, *Nature Genet.* 7:13; Lonberg *et al.*, 1994, *Nature* 368:856; and Taylor *et al.*, 1994, *Int. Immunol.* 6:579, which are hereby incorporated by reference.

The term "antibody" as used in this invention includes intact molecules as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv which are capable of binding the epitopic determinant. These antibody fragments retain some ability to selectively bind with its antigen or receptor and are defined as follows:

(1) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule can be produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain;

(2) Fab', the fragment of an antibody molecule can be obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule;

(3) $(Fab')_2$, the fragment of the antibody that can be obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; $F(ab')_2$ is a dimer of two Fab' fragments held together by two disulfide bonds;

(4) Fv, defined as a genetically engineered fragment containing the variable region of the light chain and the variable region of the heavy chain expressed as two chains; and

(5) Single chain antibody ("SCA"), defined as a genetically engineered molecule containing the variable region of the light chain, the variable region of the heavy chain, linked by a suitable polypeptide linker as a genetically fused single chain molecule.

Methods of making these fragments are known in the art. (See for example, Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1988), incorporated herein by reference). As used in this

invention, the term "epitope" means any antigenic determinant on an antigen to which the paratope of an antibody binds. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

Antibody fragments of the present invention can be prepared by proteolytic hydrolysis of the antibody or by expression in *E. coli* of DNA encoding the fragment. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. For example, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted $F(ab')_2$. This fragment can be further cleaved using a thiol reducing agent, and optionally a blocking group for the sulfhydryl groups resulting from cleavage of disulfide linkages, to produce 3.5S Fab' monovalent fragments. Alternatively, an enzymatic cleavage using pepsin produces two monovalent Fab' fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patents No. 4,036,945 and No. 4,331,647, and references contained therein. These patents are hereby incorporated in their entireties by reference. *See also* Nisonhoff *et al.*, 1960, *Arch. Biochem. Biophys.* 89:230, Porter, 1959, *Biochem. J.* 73:119; Edelman *et al.*, 1967, Methods in Enzymology, Vol. 1, page 422 (Academic Press); and Coligan *et al.* at sections 2.8.1-2.8.10 and 2.10.1-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical, or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

For example, Fv fragments comprise an association of V_H and V_L chains. This association may be noncovalent, as described in Inbar *et al.*, 1972, *Proc. Nat'l Acad. Sci. USA* 69:2659. Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde. *See, e.g.*, Sandhu, *supra*. Preferably, the Fv fragments comprise V_H and V_L chains connected by a peptide linker. These single-chain antigen binding proteins (sFv) are

prepared by constructing a structural gene comprising DNA sequences encoding the V_H and V_L domains connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell such as *E. coli*. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing sFvs are described, for example, by Whitlow *et al.*, 1991, Methods: a Companion to Methods in Enzymology, Vol. 2, page 97; Bird *et al.*, 1988, *Science* 242:423-426; Ladner *et al.*, U.S. patent No. 4,946,778; Pack *et al.*, 1993, *Bio/Technology* 11:1271-77; and Sandhu, *supra*.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells. See, for example, Larrick *et al.*, Methods: a Companion to Methods in Enzymology, Vol. 2, page 106 (1991).

Antibodies which bind to the human galectin-4 polypeptide of the invention can be prepared using an intact polypeptide or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or a peptide used to immunize an animal can be derived from translated cDNA or chemical synthesis which can be conjugated to a carrier protein, if desired. Such commonly used carriers which are chemically coupled to the peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. The coupled peptide is then used to immunize the animal (*e.g.*, a mouse, a rat, or a rabbit).

If desired, polyclonal or monoclonal antibodies can be further purified, for example, by binding to and elution from a matrix to which the polypeptide or a peptide to which the antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (See for example, Coligan, *et al.*, Unit 9, *Current Protocols in Immunology*, Wiley

Interscience, 1991, incorporated by reference).

It is also possible to use the anti-idiotype technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the "image" of the epitope bound by the first monoclonal antibody.

DETECTION OF A HUMAN GALECTIN-4 ASSOCIATED DISORDER

The antibodies and polynucleotides of the invention can be used to detect or to treat a human galectin-4-associated disorder. The term "human galectin-4-associated disorder" denotes malignant as well as non-malignant disorders, where the cells involved in the disorder differ from the surrounding tissue or from unaffected cells in their expression of human galectin-4. Malignant cells (*i.e.* cancer) develop as a result of a multistep process. The human galectin-4 polynucleotide that is an antisense molecule is useful in treating human galectin-4 disorders of the various organ systems, particularly, for example, adenocarcinomas. An "adenocarcinoma" is a carcinoma derived from glandular tissue in which the tumor cells form recognizable glandular structures. Adenocarcinomas may de classified according to the predominant pattern of cell arrangement, as papillary, alveolar, etc., or according to a particular product of the cells, as a mucinous adenocarinoma. Adenocarinomas include adenocarcinomas of the breast. Essentially, any disorder which is etiologically linked to increased expression of human galectin-4 could be considered susceptible to treatment with a human galectin-4 suppressing reagent, and any disorder which is etiologically linked to decreased expression of human galectin-4 could be considered susceptible to treatment with polynucleotides encoding human galectin-4 or the human galectin-4 polypeptide itself. One such disorder is a malignant cell proliferative disorder, for example. In addition, the invention may be used to identify or treat individuals who are "at risk" of developing a human galectin-4 associated disorder. These individuals may be identified by a method of the invention for detecting the presence or absence of human galectin-4 or by any other diagnostic

means, and or may be treated by a method of the invention, prior to the actual onset of the clinical appearance of disorder. The "clinical appearance" can be any sign or symptom of the disorder.

For purposes of the invention, an antibody or nucleic acid probe specific for human galectin-4 may be used to detect human galectin-4 polypeptide (using antibody) or polynucleotide (using nucleic acid probe) in subject samples such as biological fluids, cells, tissues, or nucleic acid. Any specimen containing a detectable amount of antigen or polynucleotide can be used. Examples of biological fluids of use with the invention are blood, serum, plasma, urine, mucous, and saliva. Tissue or cell samples can also be used with the subject invention. The samples can be obtained by many methods such as cellular aspiration, or by surgical removal of a biopsy sample. In the case of breast cancer, a tissue sample can be obtained by fine-needle aspiration, core- or cutting-needle biopsy, and excisional biopsy.

The invention provides a method for detecting a human galectin-4 associated disorder of tissue, for example, which comprises contacting an anti-human galectin-4 antibody or nucleic acid probe with a cell suspected of having a human galectin-4-associated disorder and detecting binding to the antibody or nucleic acid probe. The antibody reactive with human galectin-4 or the nucleic acid probe is preferably labeled with a compound which allows detection of binding to human galectin-4. A preferred sample of this invention is breast tissue. The level of human galectin-4 in the suspect cell can be compared with the level in a normal cell to determine whether the subject has a human galectin-4-associated cell proliferative disorder. Preferably the subject is human.

When the cell component is nucleic acid, it may be necessary to amplify the nucleic acid prior to binding with a human galectin-4 specific probe. Preferably, polymerase chain reaction (PCR) is used, however, other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleic acid sequence-based amplification (NASBA) may be used.

The antibodies of the invention can be used in any subject in which it is desirable to administer *in vitro* or *in vivo* immunodiagnosis or immunotherapy. The

antibodies of the invention are suited for use, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the antibodies in these immunoassays can be detectably labeled in various ways. Examples of types of immunoassays which can utilize antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the antibodies of the invention can be done utilizing immunoassays which are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

The antibodies of the invention can be bound to many different carriers and used to detect the presence of an antigen comprising the polypeptide of the invention. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding antibodies, or will be able to ascertain such, using routine experimentation.

There are many different labels and methods of labeling known to those of ordinary skill in the art. Examples of the types of labels which can be used in the present invention include enzymes, radioisotopes, fluorescent compounds, colloidal metals, chemiluminescent compounds, phosphorescent compounds, and bioluminescent compounds. Those of ordinary skill in the art will know of other suitable labels for binding to the antibody, or will be able to ascertain such, using routine experimentation.

Another technique which may also result in greater sensitivity consists of coupling the antibodies to low molecular weight haptens. These haptens can then be specifically detected by means of a second reaction. For example, it is common to use such haptens as biotin, which reacts with avidin, or dinitrophenyl, puridoxal, and

fluorescein, which can react with specific antihapten antibodies.

In using the monoclonal antibodies of the invention for the *in vivo* detection of antigen, the detectably labeled antibody is given a dose which is diagnostically effective. The term "diagnostically effective" means that the amount of detectably labeled monoclonal antibody is administered in sufficient quantity to enable detection of the site having the antigen comprising a polypeptide of the invention for which the monoclonal antibodies are specific.

The concentration of detectably labeled monoclonal antibody which is administered should be sufficient such that the binding to those cells having the polypeptide is detectable compared to the background. Further, it is desirable that the detectably labeled monoclonal antibody be rapidly cleared from the circulatory system in order to give the best target-to-background signal ratio.

As a rule, the dosage of detectably labeled monoclonal antibody for *in vivo* diagnosis will vary depending on such factors as age, sex, and extent of disease of the individual. Such dosages may vary, for example, depending on whether multiple injections are given, antigenic burden, and other factors known to those of skill in the art.

For *in vivo* diagnostic imaging, the type of detection instrument available is a major factor in selecting a given radioisotope. The radioisotope chosen must have a type of decay which is detectable for a given type of instrument. Still another important factor in selecting a radioisotope for *in vivo* diagnosis is that deleterious radiation with respect to the host is minimized. Ideally, a radioisotope used for *in vivo* imaging will lack a particle emission, but produce a large number of photons in the 140-250 keV range, which may readily be detected by conventional gamma cameras.

For *in vivo* diagnosis radioisotopes may be bound to immunoglobulin either directly or indirectly by using an intermediate functional group. Intermediate functional groups which often are used to bind radioisotopes which exist as metallic ions to immunoglobulins are the bifunctional chelating agents such as d-iethylenetriaminepentacetic acid (DTPA) and ethylenediaminetetraacetic acid

(EDTA) and similar molecules. Typical examples of metallic ions which can be bound to the monoclonal antibodies of the invention are ^{111}In, ^{97}Ru, ^{67}Ga, ^{68}Ga, ^{72}As, ^{89}Zr, and ^{201}Tl.

The monoclonal antibodies of the invention can also be labeled with a paramagnetic isotope for purposes of *in vivo* diagnosis, as in magnetic resonance imaging (MRI) or electron spin resonance (ESR). In general, any conventional method for visualizing diagnostic imaging can be utilized. Usually gamma and positron emitting radioisotopes are used for camera imaging and paramagnetic isotopes for MRI. Elements which are particularly useful in such techniques include ^{157}Gd, ^{55}Mn, ^{162}Dy, ^{52}Cr, and ^{56}Fe.

The monoclonal antibodies or polynucleotides of the invention can be used *in vitro* and *in vivo* to monitor the course of amelioration of a human galectin-4-associated disease in a subject. Thus, for example, by measuring the increase or decrease in the number of cells expressing antigen comprising a polypeptide of the invention or changes in the concentration of such antigen present in various body fluids, it would be possible to determine whether a particular therapeutic regimen aimed at ameliorating the human galectin-4-associated disease is effective. The term "ameliorate" denotes a lessening of the detrimental effect of the human galectin-4-associated disease in the subject receiving therapy.

The present invention identifies a nucleotide sequence that can be expressed in an altered manner as compared to expression in a normal cell, therefore it is possible to design appropriate therapeutic or diagnostic techniques directed to this sequence. Detection of elevated levels of human galectin-4 expression is accomplished by hybridization of nucleic acids isolated from a cell suspected of having a human galectin-4-associated disorder with a human galectin-4 polynucleotide of the invention. Analysis, such as Northern Blot analysis, are utilized to quantitate expression of human galectin-4. Other standard nucleic acid detection techniques will be known to those of skill in the art.

Treatment of a human galectin-4-associated disorder include modulation of human galectin-4 gene expression and human galectin-4 activity by administration

of a therapeutically effective amount of a reagent that modulates human galectin-4. The term "modulate" envisions the suppression of expression of human galectin-4 when it is over-expressed, or augmentation of human galectin-4 expression when it is under-expressed. Where a human galectin-4 associated disorder is associated with the increased expression of human galectin-4, nucleic acid sequences that interfere with human galectin-4 expression at the translational level can be used. This approach utilizes, for example, antisense nucleic acid, ribozymes, or triplex agents to block transcription or translation of a specific human galectin-4 mRNA, either by masking that mRNA with an antisense nucleic acid or triplex agent, or by cleaving it with a ribozyme. Such disorders include breast cancer, for example.

Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (Weintraub, 1990, *Scientific American*, 262:40). In the cell, the antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule. The antisense nucleic acids interfere with the translation of the mRNA, since the cell will not translate a mRNA that is double-stranded. Antisense oligomers of about 15 nucleotides are preferred, since they are easily synthesized and are less likely to cause problems than larger molecules when introduced into the target human galectin-4-producing cell. The use of antisense methods to inhibit the *in vitro* translation of genes is well known in the art (Marcus-Sakura, 1988, *Anal.Biochem.*, 172:289).

Use of an oligonucleotide to stall transcription is known as the triplex strategy since the oligomer winds around double-helical DNA, forming a three-strand helix. Therefore, these triplex compounds can be designed to recognize a unique site on a chosen gene (Maher, *et al.*, 1991, *Antisense Res. and Dev.*, 1(3):227; Helene, C., 1991, *Anticancer Drug Design*, 6(6):569).

Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences which encode these RNAs, it is possible to engineer molecules that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, 1988, *J.Amer.Med. Assn.*,

260:3030). A major advantage of this approach is that, because they are sequence-specific, only mRNAs with particular sequences are inactivated.

There are two basic types of ribozymes namely, *tetrahymena*-type (Hasselhoff, 1988, *Nature*, 334:585) and "hammerhead"-type. *Tetrahymena*-type ribozymes recognize sequences which are four bases in length, while "hammerhead"-type ribozymes recognize base sequences 11-18 bases in length. The longer the recognition sequence, the greater the likelihood that the sequence will occur exclusively in the target mRNA species. Consequently, hammerhead-type ribozymes are preferable to *tetrahymena*-type ribozymes for inactivating a specific mRNA species and 18-based recognition sequences are preferable to shorter recognition sequences.

The present invention also provides gene therapy for the treatment of cell proliferative or other disorders which are associated with by human galectin-4 protein. Such therapy would achieve its therapeutic effect by introduction of the human galectin-4 antisense polynucleotide into cells having the disorder. Delivery of antisense human galectin-4 polynucleotide can be achieved using a recombinant expression vector such as a chimeric virus or a colloidal dispersion system. Especially preferred for therapeutic delivery of antisense sequences is the use of targeted liposomes.

Various viral vectors which can be utilized for gene therapy as taught herein include adenovirus, herpes virus, vaccinia, or, preferably, an RNA virus such as a retrovirus. Preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of retroviral vectors in which a single foreign gene can be inserted include, but are not limited to: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), and Rous Sarcoma Virus (RSV). Preferably, when the subject is a human, a vector such as the gibbon ape leukemia virus (GaLV) is utilized. A number of additional retroviral vectors can incorporate multiple genes. All of these vectors can transfer or incorporate a gene for a selectable marker so that transduced cells can be identified and generated. By inserting a human galectin-4 sequence of interest into the viral

vector, along with another gene which encodes the ligand for a receptor on a specific target cell, for example, the vector is now target specific. Retroviral vectors can be made target specific by attaching, for example, a sugar, a glycolipid, or a protein. Preferred targeting is accomplished by using an antibody to target the retroviral vector. Those of skill in the art will know of, or can readily ascertain without undue experimentation, specific polynucleotide sequences which can be inserted into the retroviral genome or attached to a viral envelope to allow target specific delivery of the retroviral vector containing the human galectin-4 antisense polynucleotide.

Since recombinant retroviruses are defective, they require assistance in order to produce infectious vector particles. This assistance can be provided, for example, by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR. These plasmids are missing a nucleotide sequence which enables the packaging mechanism to recognize an RNA transcript for encapsidation. Helper cell lines which have deletions of the packaging signal include, but are not limited to Q2, PA317 and PA12, for example. These cell lines produce empty virions, since no genome is packaged. If a retroviral vector is introduced into such cells in which the packaging signal is intact, but the structural genes are replaced by other genes of interest, the vector can be packaged and vector virion produced.

Alternatively, NIH 3T3 or other tissue culture cells can be directly transfected with plasmids encoding the retroviral structural genes *gag, pol* and *env*, by conventional calcium phosphate transfection. These cells are then transfected with the vector plasmid containing the genes of interest. The resulting cells release the retroviral vector into the culture medium.

Another targeted delivery system for human galectin-4 antisense polynucleotides is a colloidal dispersion system. Colloidal dispersion systems include macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. The preferred colloidal system of this invention is a liposome. Liposomes are artificial membrane vesicles which are useful as delivery vehicles *in vitro* and *in vivo*.

It has been shown that large unilamellar vesicles (LUV), which range in size from 0.2-4.0 μm can encapsulate a substantial percentage of an aqueous buffer containing large macromolecules. RNA, DNA and intact virions can be encapsulated within the aqueous interior and be delivered to cells in a biologically active form (Fraley, *et al.*, 1981, *Trends Biochem. Sci.*, 6:77,). In addition to mammalian cells, liposomes have been used for delivery of polynucleotides in plant, yeast and bacterial cells. In order for a liposome to be an efficient gene transfer vehicle, the following characteristics should be present: (1) encapsulation of the genes of interest at high efficiency while not compromising their biological activity; (2) preferential and substantial binding to a target cell in comparison to non-target cells; (3) delivery of the aqueous contents of the vesicle to the target cell cytoplasm at high efficiency; and (4) accurate and effective expression of genetic information (Mannino, *et al.*, 1988, *B-iotechniques*, 6:682).

The composition of the liposome is usually a combination of p-hospholipids, particularly high-phase-transition-temperature phospholipids, usually in combination with steroids, especially cholesterol. Other phospholipids or other lipids may also be used. The physical characteristics of liposomes depend on pH, ionic strength, and the presence of divalent cations.

Examples of lipids useful in liposome production include phosphatidyl compounds, such as phosphatidylglycerol, phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, sphingolipids, cerebrosides, and gangliosides. Particularly useful are d-iacylphosphatidylglycerols, where the lipid moiety contains from 14-18 carbon atoms, particularly from 16-18 carbon atoms, and is saturated. Illustrative phospholipids include egg phosphatidylcholine, dipalmitoylphosphatidylcholine and distearoylphosphatidylcholine.
The targeting of liposomes can be classified based on anatomical and mechanistic factors. Anatomical classification is based on the level of selectivity, for example, organ-specific, cell-specific, and organelle-specific. Mechanistic targeting can be distinguished based upon whether it is passive or active. Passive targeting utilizes the natural tendency of liposomes to distribute to cells of the reticulo-endothelial system

(RES) in organs which contain sinusoidal capillaries. Active targeting, on the other hand, involves alteration of the liposome by coupling the liposome to a specific ligand such as a monoclonal antibody, sugar, glycolipid, or protein, or by changing the composition or size of the liposome in order to achieve targeting to organs and cell types other than the naturally occurring sites of localization.

The surface of the targeted delivery system may be modified in a variety of ways. In the case of a liposomal targeted delivery system, lipid groups can be incorporated into the lipid bilayer of the liposome in order to maintain the targeting ligand in stable association with the liposomal bilayer. Various linking groups can be used for joining the lipid chains to the targeting ligand.

This invention involves administering to a subject a therapeutically effective dose of a pharmaceutical composition containing the compounds of the present invention and a pharmaceutically acceptable carrier. "Administering" the pharmaceutical composition of the present invention may be accomplished by any means known to the skilled artisan. By "subject" is meant any mammal, preferably a human.

The pharmaceutical compositions are preferably prepared and administered in dose units. Solid dose units are tablets, capsules and suppositories. For treatment of a patient, depending on activity of the compound, manner of administration, nature and severity of the disorder, age and body weight of the patient, different daily doses are necessary. Under certain circumstances, however, higher or lower daily doses may be appropriate. The administration of the daily dose can be carried out both by single administration in the form of an individual dose unit or else several smaller dose units and also by multiple administration of subdivided doses at specific intervals.

The pharmaceutical compositions according to the invention are in general administered topically, intravenously, orally or parenterally or as implants, but even rectal use is possible in principle. Suitable solid or liquid pharmaceutical preparation forms are, for example, granules, powders, tablets, coated tablets, (micro)capsules, suppositories, syrups, emulsions, suspensions, creams, aerosols, drops or injectable

solution in ampule form and also preparations with protracted release of active compounds, in whose preparation excipients and additives and/or auxiliaries such as disintegrants, binders, coating agents, swelling agents, lubricants, flavorings, sweeteners or solubilizers are customarily used as described above. The pharmaceutical compositions are suitable for use in a variety of drug delivery systems. For a brief review of present methods for drug delivery, *see* Langer, 1990, *Science*, 249:1527-1533, which is incorporated herein by reference.

The pharmaceutical compositions according to the invention may be administered locally or systemically. By "therapeutically effective dose" is meant the quantity of a compound according to the invention necessary to prevent, to cure or at least partially arrest the symptoms of the disorder and its complications. Amounts effective for this use will, of course, depend on the severity of the disease and the weight and general state of the patient. Typically, dosages used *in vitro* may provide useful guidance in the amounts useful for *in situ* administration of the pharmaceutical composition, and animal models may be used to determine effective dosages for treatment of particular disorders. Various considerations are described, *e.g.*, in Gilman *et al.* (eds.) (1990) Goodman And Gilman's: The Pharmacological Bases of Therapeutics, 8th ed., Pergamon Press; and Remington's Pharmaceutical Sciences, 17th ed. (1990), Mack Publishing Co., Easton, Pa., each of which is herein incorporated by reference.

The identification of a human galectin-4 provides a useful tool for diagnosis, prognosis and therapeutic strategies associated with human galectin-4 mediated disorders. Measurement of human galectin-4 levels using anti-human galectin-4 antibodies is a useful diagnostic for following the progression or recovery from breast cancer, or for determining the presence of metastases in a subject.

KITS

The materials for use in the assay of the invention are ideally suited for the preparation of a kit. Such a kit may comprise a carrier means being compartmentalized to receive in close confinement one or more container means such

as vials, tubes, and the like, each of the container means comprising one of the separate elements to be used in the method. One of the container means may comprise a probe which is or can be detectably labeled. Such probe may be an antibody or nucleotide specific for human galectin-4, or fragments thereof. For example, oligonucleotide probes of the present invention can be included in a kit and used for examining the presence of human galectin-4, as well as the quantitative (relative) degree of binding of the probe for determining the occurrence of specific strongly binding (hybridizing) sequences, thus indicating the likelihood for an subject having or predisposed to a disorder associated with human galectin-4.

The kit may also contain a container comprising a reporter-means, such as a biotin-binding protein, such as avidin or streptavidin, bound to a reporter molecule, such as an enzymatic, fluorescent, or radionucleotide label to identify the detectably labeled oligonucleotide probe.

Where the kit utilizes nucleic acid hybridization to detect the target nucleic acid, the kit may also have containers containing nucleotide(s) for amplification of the target nucleic acid sequence. When it is desirable to amplify the human galectin-4 target sequence, this can be accomplished using oligonucleotide(s) that are primers for amplification. These oligonucleotide primers are based upon identification of the flanking regions contiguous with the target nucleotide sequence.

The kit may also contain a container containing antibodies which bind to human galectin-4, or fragments thereof. Such antibodies can be used to distinguish the presence of human galectin-4 or the level of expression of human galectin-4 in a specimen. Where the kit utilizes antibodies to detect human galectin-4, these antibodies may be directly labeled. The kit may also contain a container containing a reporter means, such as aviden or steptavin, bound to a reporter molecule such as an enzymatic, fluorescent, or radionucleotide label to identify the directly labeled antibody. Alternatively, the kit can utilizes antibodies that bind human galectin-4 that are unlabeled. The kit may then also contain a container containing a second antibody which binds to the antibody specific for human galectin-4. The second antibody can be directly labeled. The kit may also a container containing a reporter

means, such as aviden or steptavin, bound to a reporter molecule such as an enzymatic, fluorescent, or radionucleotide label to identify the directly labeled second antibody.

SCREENING FOR COMPOUNDS WHICH AFFECT EXPRESSION

In another embodiment, the invention provides a method for identifying a compound which affects human galectin-4 expression. The method includes incubating compounds and a human galectin-4 positive cell under conditions sufficient to allow the components to interact and measuring the level of human galectin-4 in the cells with a reagent of the invention such as an antibody which binds human galectin-4 polynucleotide or a polynucleotide that specifically binds human galectin-4 nucleic acid. The compounds which affect human galectin-4 include peptides, polypeptides, chemical compounds and biological agents.

"Incubating" includes conditions which allow contact between the test compound and human galectin-4. "Contacting" includes in solution and solid phase. The test compound may also be a combinatorial library for screening a plurality of compounds. Compounds identified in the method of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution of after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence, such as PCR, oligomer restriction (Saiki *et al.*, 1985, *Bio/Technology*, 3:1008-1012), allele-specific oligonucleotide (ASO) probe analysis (Conner *et al.*, 1983, *Proc. Natl. Acad. Sci. USA*, 80:278), oligonucleotide ligation assays (OLAs) (Landegren *et al.*, 1988, *Science*, 241:1077), and the like. Molecular techniques for DNA analysis have been reviewed (Landegren *et al.*, 1988, *Science*, 242:229-237).

TRANSGENIC ANIMALS

In another embodiment, the present invention relates to transgenic animals having cells that express human galectin-4. Such transgenic animals represent a model system for the study of human galectin-4 related disorders and the study of human galectin-4 based therapeutics.

The term "animal" here denotes all mammalian species except human. It also includes an individual animal in all stages of development, including embryonic and fetal stages. Farm animals (pigs, goats, sheep, cows, horses, rabbits and the like), rodents (such as mice), and domestic pets (for example, cats and dogs) are included within the scope of the present invention.

A "transgenic" animal is any animal containing cells that bear genetic information received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by microinjection or infection with recombinant virus. "Transgenic" in the present context does not encompass classical crossbreeding or *in vitro* fertilization, but rather denotes animals in which one or more cells receive a recombinant DNA molecule. Although it is highly preferred that this molecule be integrated within the animal's chromosomes, the present invention also contemplates the use of extrachromosomally replicating DNA sequences, such as might be engineered into yeast artificial chromosomes.

The term "transgenic animal" also includes a "germ cell line" transgenic animal. A germ cell line transgenic animal is a transgenic animal in which the genetic information has been taken up and incorporated into a germ line cell, therefore conferring the ability to transfer the information to offspring. If such offspring in fact possess some or all of that information, then they, too, are transgenic animals.

It is highly preferred that the transgenic animals of the present invention be produced by introducing into single cell embryos DNA encoding human galectin-4, in a manner such that the polynucleotides are stably integrated into the DNA of germ line cells of the mature animal and inherited in normal mendelian fashion. Advances in technologies for embryo micromanipulation now permit introduction of heterologous DNA into fertilized mammalian ova. For instance,

totipotent or pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In a preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo.

In a most preferred method the appropriate DNAs are coinjected into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals. These techniques are well known. For instance, reviews of standard laboratory procedures for microinjection of heterologous DNAs into mammalian (mouse, pig, rabbit, sheep, goat, cow) fertilized ova include: Hogan *et al.*, Manipulating the Mouse Embryo (Cold Spring Harbor Press 1986); Krimpenfort *et al.*, 1991, *Bio/Technology* 9:86; Palmiter *et al.*, 1985, *Cell* 41:343; Kraemer *et al.*, Genetic Manipulation of the Early Mammalian Embryo (Cold Spring Harbor Laboratory Press 1985); Hammer *et al.*, 1985, *Nature*, 315:680; Purcel *et al.*, 1986, *Science*, 244:1281; Wagner *et al.*, U.S. patent No. 5,175,385; Krimpenfort *et al.*, U.S. patent No. 5,175,384, the respective contents of which are incorporated by reference.

The cDNA that encodes human galectin-4 can be fused in proper reading frame under the transcriptional and translational control of a vector to produce a genetic construct that is then amplified, for example, by preparation in a bacterial vector, according to conventional methods. See, for example, the standard work: Sambrook *et al.*, Molecular Cloning: a Laboratory Manual (Cold Spring Harbor Press 1989), the contents of which are incorporated by reference. The amplified construct is thereafter excised from the vector and purified for use in producing transgenic animals.

The term "transgenic" as used herein additionally includes any organism whose genome has been altered by *in vitro* manipulation of the early embryo or fertilized egg or by any transgenic technology to induce a specific gene knockout. The term "gene knockout@ as used herein, refers to the targeted disruption of a gene

in vivo with complete loss of function that has been achieved by any transgenic technology familiar to those in the art. In one embodiment, transgenic animals having gene knockouts are those in which the target gene has been rendered nonfunctional by an insertion targeted to the gene to be rendered non-functional by homologous recombination. As used herein, the term "transgenic" includes any transgenic technology familiar to those in the art which can produce an organism carrying an introduced transgene or one in which an endogenous gene has been rendered non-functional or knocked out.

The transgene to be used in the practice of the subject invention is a DNA sequence comprising a modified human galectin-4 coding sequence. In a preferred embodiment, the human galectin-4 gene is disrupted by homologous targeting in embryonic stem cells. For example, the entire human galectin-4 gene may be deleted. Optionally, the human galectin-4 disruption or deletion may be accompanied by insertion of or replacement with other DNA sequences, such as a non-functional human galectin-4 sequence. In other embodiments, the transgene comprises DNA antisense to the coding sequence for human galectin-4. In another embodiment, the transgene comprises DNA encoding an antibody or receptor peptide sequence which is able to bind to human galectin-4. Where appropriate, DNA sequences that encode proteins having human galectin-4 activity but differ in nucleic acid sequence due to the degeneracy of the genetic code may also be used herein, as may truncated forms, allelic variants and interspecies homologues.

The following examples are intended to illustrate but not limit the invention. While they are typical of those that might be used, other procedures known to those skilled in the art may alternatively be used.

EXAMPLES

Example 1
Purification and sequence of human galectin-4

A. Identification of galectin-4 and galectin-3 in T84 cells.

To study the localization and function of galectins in a cell culture model of intestinal epithelium, we looked for these lactose binding proteins in T84 colon carcinoma cells by affinity chromatography of cell extracts on lactosyl-Sepharose. T84 cells (passages 54-70) were grown following the procedure of Dharmsathaphorn *et al.* (1984), in DME H-16/F-12 medium (1:1) containing 5% newborn calf serum, 5% FCS, penicillin (100 units/ml), streptomycin (100μg/ml), 15mM HEPES pH 7.3, 17.5mM glucose and 2.5mM glutamine, in a humidified atmosphere of 95% air and 5% CO2 at 37°C (seeding density of 2.5 X 105 cells/cm2).

The cells were labeled *in vivo* with [35S]-Met/Cys. An [^{35}S] protein labeling mix EXPRE35S35S (>1000Ci/mmol) was purchased from Du Pont NEN (Boston, MA). For labeling *in vivo* with [35S]-Met/Cys, T84 cells were grown on 10 cm diameter plastic dishes as described above. Endogenous methionine/cysteine was depleted before labeling, by washing cells with methionine-free MEM, and incubating in "depletion medium" (methionine-free MEM, 5% dialyzed FCS, 10 mM HEPES pH 7.3), for 1 hour at 37°C. For labeling, the medium was replaced with 5 ml of the same medium containing 1 mCi of [35S]-methionine/cysteine. After 12 hours at 37°C the cells were washed extensively, first in prewarmed, then in ice-cold PBS (5.4 mM sodium/potassium phosphate, 135 mM NaCl, pH 7.4), and lysed in 1.5ml/plate of ice-cold "lysis buffer": 1 mM EGTA, 1.5 mM MgCl2, 5 mM b-mercaptoethanol, 2.5% Triton X-100 (v/v) in PBS, containing protease inhibitors (3mM PMSF, 10 μg/ml leupeptin, 10 μg/ml aprotinin). The lysate was centrifuged at 1000 x g for 15 min. at 4°C, and the supernatant was used for purification of galectins by affinity chromatography.

The supernatant of the cell lysate was passed over a lactosyl-Sepharose

column (prepared as described by Leffler *et al.*, 1989) equilibrated with the lysis buffer without protease inhibitors. Unbound material was washed off the column with buffer containing 135 mM NaCl, 20 mM Tris-HCl pH 7.5 (TBS), and 5 mM b-mercaptoethanol. Lectin was eluted with buffer containing 150 mM lactose, 50 mM NaCl and 10 mM Tris-HCl pH 7.5. Protein concentration was measured using BioRad protein assay (BioRad, Richmond, CA). For SDS-PAGE the MiniProtean unit (BioRad) was used under standard conditions.

All samples were precleared twice with Sepharose CL-2B (Pharmacia ,Piscataway, NJ) before addition of the primary antibody. The washed precipitates were boiled in SDS-PAGE sample buffer and separated on a 12% polyacrylamide SDS gel. Proteins were localized by Coomassie blue staining followed by fluorography at -70°C with Kodak XAR-5 film and intensifying screens. The ratio of galectin-3 to galectin-4 was estimated by densitometric scanning. We found that T84 cells express high concentrations of a 36 kDa and a 29 kDa galectin (Fig. 1, lanes a and b). The former reacted with anti-galectin-4 on Western blots (Fig. 1, lane c) and the latter was immunoprecipitated by monoclonal anti-galectin-3 antibodies (Fig. 1, lane d). For the immunoprecipitation reactions, galectin-3 was immunoprecipitated from both T84 cell lysate or from galectin-containing fractions eluted from lactosyl-Sepharose using anti-galectin-3 and Protein G-Sepharose (Zymed, South San Francisco, CA) as described (Huflejt *et al.*, 1993).

The relative combined amount of galectin-3 and -4 in T84 cells was estimated as the amount of 35S-methionine/cysteine incorporated in the galectins compared to total TCA preciptable protein from 2.5% Triton X-100 cell lysates. The combined radioactivity incorporated into the two galectins represents 1.4 % of the TCA-precipitated radioactivity in early confluent cells but declined to about 0.5% in aging cells. Densitometry of SDS-PAGE of the purified galectin mixture showed that galectin-4 accounts for 38-60% of the total galectin content.

The molar concentration of the galectins in T84 ells was calculated based on the absolute yield of lactosyl-Sepharose purified galectins from a given area of confluent monolayer. Thus, about 0.5 mg galectin was obtained from ten culture

plates (10 cm diameter), corresponding to a cell monolayer area of 800 cm2. Since the cells are on average 25 mm high (Fig. 3), this corresponds to a cell volume of 2 ml. Hence the combined concentration of the galectins is 0.25 mg/ml or about 8 mM. Based on the relative amounts of each galectin estimated by densitometry the average concentration of galectin-3 and -4 in T84 cells is about 5 and 3 mM, respectively.

<u>B.</u> <u>cDNA cloning</u>

The expression of galectin-4 in T84 cells was confirmed by cDNA cloning and sequencing (Fig. 2A). Unless specified otherwise, all manipulations of nucleic acids such as restriction, ligation, transformation, gel electrophoresis, blotting, gel elution, radiolabeling, and preparation of buffers were done using standard protocols (Sambrook *et al.*, 1990). Samples were sequenced using Sequenase v2.0 kit (USB, Cleveland, OH). The reported sequence was confirmed on both strands.

Human genomic DNA was amplified with convergent oligonucleotide primer pairs complementary to several different parts of the rat galectin-4 gene (Oda *et al.*, 1993), by PCR using 1 μg of human genomic DNA (Clontech, Palo Alto, CA) per 50 μl reaction. Reactions were carried out under conditions described by Gitt. *et al.*, (1995), for 45 cycles: 40 sec at 95°C, 1 min. at 55°C, and 4 min. at 72°C. PCR using primers 5'CTGCCATGGCGGACCCCCGATCTTCAA3' (SEQ ID NO:3) and 5'ATGATGGTTCTTCGGGCTG3' (SEQ ID NO:4) yielded a major product which was cloned into the pCR1000 vector (Invitrogen, San Diego, CA) according to the manufacturer's instructions. This product was a gene fragment containing an intron flanked by two exon fragments. A human-specific primer, HL36A, 5'TGAGCCCTCCTTGCAGCC3' (SEQ ID NO:5) was designed based on the exon sequence, to be used in PCR of reverse transcribed RNA.

Total RNA was isolated from confluent T84 cells using RNAzol (Tel-Test, Friendswood, TX) according to the manufacturer's protocol. The RNA was reverse-transcribed with the C35 primer, 5'GACTCGAGTCGACATCGATTTTTTTTTTTTTTTTT3' (SEQ ID NO:6) using MMLV Reverse Transcriptase. The 20 μl solution was then diluted to 1 ml with 10

mM Tris-HCl pH 7.6, 1 mM EDTA, and 5 μl of this diluted solution was used in PCR with the primers HL36A and 1 (5'TTCTATGAATACGGGCACCGG3') (SEQ ID NO:7): 5 cycles of annealing at 50°C for 1 min. and 40 cycles of annealing at 60°C for 1 min. Extension at 72°C was for 2 min. in all cases. One pure band was obtained, cloned into pCR1000 and sequenced using vector-specific primers. Based on this sequence, another human-specific primer, HL36B, 5'TACCCTGGTCCCGGACATTG3' (SEQ ID NO:8), was synthesized, and the next PCR was performed using HL36B and C19 (a primer containing the sequence of primer C35 without the oligoT stretch). Annealing was at 56°C for 1 min. and polymerization was at 72°C for 1.5 min. A single band was obtained and ligated into pCR1000, yielding a clone which was used as a probe to screen a cDNA library prepared from T84 cells (a generous gift from Dr. J.R. Riordan, Sick Children's Hospital, Toronto). Positive clones were purified by rescreening. One positive clone was found to contain all previously identified sequence and additional 5' sequence.

The amino acid sequence of the encoded T84 cell protein showed about 80% sequence identity with rat and porcine galectin-4 within the carbohydrate-binding domains, and about 50% sequence identity within the link region, and contained all residues that are typically conserved in galectins and known to be associated with carbohydrate binding activity (Fig. 2B). From these data, we conclude that the 36 kDa lactose-binding protein of T84 cells is human galectin-4.

C. <u>Immunocytochemistry of galectins</u>

The antibodies used in these assays were as follows: a rat monoclonal anti-galectin-3 (anti-Mac-2; Ho and Springer, 1982; Cherayil *et al.*, 1989) was used as described (Huflejt et al., 1993). Rabbit anti-galectin-4 serum was raised against the C-terminal domain of rat intestinal galectin-4 as described (Oda *et al.*, 1993). A monoclonal anti-cytokeratin antibody 7D3 was a generous gift from Dr. Caroline Damsky, Univ. of California, San Francisco (Damsky *et al.*, 1992). A mouse monoclonal antibody (IgG2a) against the intracellular domain of human E-cadherin was purchased from Transduction Laboratories (Lexington, KY). Fluorescein-

conjugated rabbit anti-rat antibodies, biotinylated sheep anti-rat antibodies, streptavidin-conjugated fluorescein, and streptavidin-conjugated Texas Red were purchased from Amersham (Arlington Heights, IL). Biotinylated goat anti-rabbit antibodies, Vectastain kit, and Vectashield were from Vector (So. San Francisco, CA).

Cells grown on uncoated cover slips were washed with PBS and fixed either with 4% paraformaldehyde-periodate-lysine (PLP; Brown and Farquhar, 1989), or with 4% paraformaldehyde in PBS for 12 h at room temperature, and washed repeatedly with PBS. The cells were permeabilized with 0.025% saponin in PBS for 10 min. at 37°C. Nonspecific antibody binding was blocked by incubation of fixed cells in PBS containing 6.6 mg/ml fish skin gelatin (Sigma, St. Louis, MO), (PBS-FSG), and cells were incubated with primary antibody in PBS-FSG for 1h at room temperature, washed 3 X 10 min. in PBS-FSG, incubated with fluorochrome- or biotin-conjugated secondary antibody in PBS-FSG for 30 min., washed, and incubated with streptavidin-conjugated second fluorochrome in PBS for 30 min. Samples were then thoroughly washed in PBS, and after a final rinse in double distilled H_2O mounted in Vectashield (Vector Laboratories, Burlingame, CA), and analyzed in a laser scanning confocal module (BioRad MRC600 confocal head attached to a Nikon Optiphot microscope with a x63 1.4 N.A. objective, 0.5s/scan with Kalman filter, 0.3 μm/step.

The distribution of the two galectins in T84 cells was examined by immunocytochemistry and confocal microscopy. However, the immunocytochemical detection of the galectins was complicated by an unexpected difficulty in galectin fixation. Application of standard fixation and permeabilization conditions resulted in very weak staining of both galectins, suggesting that the galectins, which are very abundant in T84 cells, were lost due to poor fixation. Indeed, both galectins were found in the solutions used for cell permeabilization, and could be purified on lactosyl-Sepharose showing that they retain activity. It was particularly notable that active galectins could be identified that way even in methanol and acetone solutions (lyophilized and resolubilized in lysis buffer) collected after cell "fixation" for 5 min.

at -20°C. Therefore, several permutations of different fixation and permeabilization procedures were tested that included prolonged fixation (twelve hours with 4% paraformaldehyde-lysine-sodium periodate, PLP, or with 4% paraformaldehyde), followed by mild brief permeabilization with either saponin (0.025% in PBS, 10 min.) to visualize intracellular galectins in confluent cells, or standard PLP-fixation without permeabilization to visualize membrane associated galectins in EGTA treated confluent and in subconfluent cells.

Figure 3a presents the localization of galectin-4 (red) and galectin-3 (green) in 25 μm tall, confluent T84 cells as XZ sections after prolonged PLP-fixation and brief saponin permeabilization. Galectin-4 is seen mainly as a 1-2 mm thick layer near the basal membrane. Galectin-3 is mainly found in subapical accumulations revealed more clearly after fixation with paraformaldehyde instead of PLP (Fig. 3b). XY sections confirmed the apical-basal polarity of galectin distribution, and also revealed significant amounts of diffusely distributed galectins (Fig. 4). Thus, panel a of Fig. 4 recorded at the level of the apical membrane (24 mm above the substratum) shows mainly galectin-3, whereas panel c recorded 8 mm above the substratum shows mainly galectin-4.

The XY sections also revealed particularly dense accumulations of the two galectins with a distinct morphology at the apical membrane (Fig. 4a, arrowheads), which were not readily seen in the XZ sections. Higher magnification of these detergent-resistant apical accumulations of galectins (Fig. 4d) demonstrates their very defined shape and composition. Galectin-4 is concentrated into circular accumulations with a large adjacent zone of galectin-3 either in the cell center (large arrow head) or at the site of cell-cell contacts (small arrow head) at the apical membrane (Fig. 4d). There is usually one organized accumulation of galectins in each confluent cell, and such formations have not been found in nonconfluent cells.

The galectin-4 localization at the adherens junctions in porcine tongue squamous epithelium (Chiu *et al.*, 1992, 1994) was not observed in the permeabilized confluent T84 monolayers. Since this might be due to the difficulties of fixing galectin-4 to its neighboring proteins resulting in its loss from permeabilized cells, as

discussed above, we used an alternative approach to look for lateral localization of galectin-4 without permeabilization. To gain access to the extracellular adherens junction area in confluent T84 cells without permeabilization, the tight junction complexes were uncoupled by brief removal of Ca2+ with EGTA, followed by fixation with PLP. Thus, to remove extracellular Ca2+, the cells were incubated in growth medium containing 4 mM EGTA pH. 7.4, for 15 or 30 min. at 37°C (Kartenbeck *et al.*, 1991), rinsed with Ca2+, Mg2+-free PBS, fixed and processed as above. These EGTA-treated confluent monolayers were double-immunolabeled for galectin-4 (red), and cytokeratin (green, Fig. 5). Cytokeratin immunostaining was used here as an indicator of the ability of antibodies to gain access to the interior of nonpermeabilized T84 cells; indeed, the anti-cytokeratin antibodies stained a layer about 2 mm inside the cell perimeter, outlining each cell, indicating that antibodies can penetrate nonpermeabilized T84 cells.

In the cells treated with EGTA for 15 min., galectin-4 is mainly observed as dramatic red patches (Fig. 5a, arrowhead) covering a large (8-16 μm long) lateral membrane area of cells next to a site of breakage in the monolayer. Following removal of Ca2+ with EGTA for 30 min. (Fig. 5b), cell shapes were much less regular and there were many more patches of galectin-4 covering the lateral membranes (Fig. 5b, arrowheads), again mostly next to a site of breakage in the monolayer. There was little if any colocalization of cytokeratin with galectin-4. Instead, in these conditions galectin-4 seemed to be localized more toward the cell periphery than cytokeratin, or extracellularly. These characteristic lateral galectin-4 patches were never seen in confluent monolayers that had not been treated with EGTA, but persisted upon permeabilization.

Since in porcine oral squamous epithelium galectin-4 colocalizes with E-cadherin (Chiu *et al.*, 1992), we compared the localization of galectin-4 (red) and the intracellular domain of E-cadherin (green, Figs. 6a-d) in EGTA-treated confluent T84 cells. After 30 min. of EGTA treatment, the main galectin-4 staining was found near the apical membrane (XY plane 24 μm above glass level, Fig. 6a), concentrated in brightly staining crescent-shaped aggregations. These crescents are found at sites

of the characteristic apical rounding and separation from neighboring cells, and correspond to the lateral galectin-4 patches shown in Figs. 5a and b. Less intense galectin-4 staining was found in form of similar crescents 8 and 16 mm below the apical membrane, where most E-cadherin was found in the form of intracellular vesicular staining that did not colocalize with galectin-4 (Figs. 6b and c, arrowheads).

It is notable that the galectin-4 staining seen near the basal membrane and diffusely cytosolic in permeabilized cells not treated with EGTA (Fig. 3-4) was diminished after 15 minutes EGTA treatment (Fig. 5a) and not seen in the cells treated for 30 minutes with EGTA even if the anti-cytokeratin and anti-E-cadherin clearly had access to these sites. In addition this basal and cytosolic staining was also absent in EGTA-treated and saponin permeabilized cells (not shown) whereas the lateral galectin-4 patches were present as mentioned above. Hence, it is possible that detergent-resistant galectin-4-rich patches at the lateral membrane are the result of an EGTA-induced accumulation of galectin-4 at this site. Alternatively, galectin-4 at this site is exposed when cell-cell junctions are broken, and the galectin-4 at other sites is lost due to independent mechanisms. Fig. 6 shows that some galectin-4 colocalizes with E-cadherin remaining at the lateral site typical of adherens junctions, but not with most of the E-cadherin that appears to have been internalized after calcium depletion (as observed in MDBK cells, Kartenbeck *et al.*, 1991).

The distribution of galectin-4 and galectin-3 in T84 cell cultures at different stages after seeding was examined by immunocytochemistry of fixed but not permeabilized cells. For these experiments, 2-3 day old subconfluent cells were dissociated by brief and mild trypsin (0.05%) and low EDTA (0.02%) treatment for 1 min. at 37°C. These brief and gentle conditions were chosen, since dissociation with higher concentration of EDTA or EGTA results in a decreased rate of cell attachment, and great loss of observable galectin-4. The cells were examined at 2 hours, 12 hours, 24 hours and 48 hours after seeding.

An XZ-section through a typical cell cluster two hours after seeding is shown in Fig 7a. In this cluster two cells are attaching to the substrate, and the top cell has not yet attached. A horizontal optical section through the top cell (Fig. 7b) shows

a remarkable concentration of galectin-4 within one hemisphere of the cell periphery, whereas galectin-3 is evenly distributed around the cell periphery. This pattern was typical for rounded and not yet attaching cells. It is possible that this polarization of galectin-4 reflects the organization of the periphery of basal and apical cytoplasm (compare with Figs. 3 and 4), retained by the cells after dissociation from the polarized monolayer.

In the two attached cells seen at the bottom of the cluster in Fig. 7a, galectin-4 is concentrated at the points of cell-substrate contact (arrow heads), whereas galectin-3 is found along a large part of the cell membrane. In 12 hour- and 24 hour-old colonies (Fig. 7c and d), galectin-4 is again concentrated to small areas of the cell periphery at the same pole of pairs of newly divided cells (Fig. 7c), and at newly formed substrate contact sites at the colony periphery (Fig. 7d, small arrowhead). In contrast, galectin-3 is more diffusely distributed in the peripheral cytoplasm, including the vicinity of the membrane separating the newly divided cells. The galectin-4-positive patches that are often noticed at the surface of the glass (large arrowhead in Fig. 7d) are probably the remnants of attachment sites of cells dislodged during washes. The thin red line seen at the level of the glass in areas not covered by cells (Fig. 7a) may also be due to adsorbed galectin-4 released from disrupted cells.

In 48 hour cultures, large colonies have formed with cells rapidly proliferating and establishing first contacts with the substrate at the edge, and more differentiated cells in the interior. Fig. 8 shows an overview of the galectin localization in such cultures revealed with immunoperoxidase instead of immunofluorescence to permit covisualization of the underlying cell and colony morphology. It is clear that both galectin-3 (panels a, c and e) and galectin-4 (panels b, d, f) are found in most lamellipodia at the edge of these fixed but not permeabilized cell colonies. In saponin- or methanol-permeabilized subconfluent cells, both galectins are essentially absent from lamellipodia (not shown), suggesting that their association with other cellular components at this location is detergent-sensitive, in contrast to the galectin-4 seen at the lateral membrane of calcium-depleted cells shown in Figs. 5a and b.

Double immunostaining and confocal microscopy of the 48 hour cultures showed that both galectins are found within the same lamellipodia, but in clearly distinct regions (Fig. 7e-f): galectin-4 is always concentrated in the leading edge, whereas galectin-3 is found distributed more diffusely in the more proximal part. This distribution is preserved both in the compact, pointed lamellipodia typical for serum starved cells (Fig. 7e), and in the wide flat (1-2 μm tall, Fig. 7f inset) but delicate and webby outgrowth areas seen 15 min. after serum stimulation. Hence, the distribution of galectins defines three areas of lamellipodia: a leading edge occupied exclusively by galectin-4, the central area where the two galectins colocalize, and a proximal region rich in galectin-3.

In conclusion, in subconfluent T84 cells, galectin-4 is found in attachment sites of newly seeded cells, and at the leading edge of lamellipodia. In contrast, galectin-3 is distributed along most of the cell periphery of these cells, and is more concentrated in the posterior part of lamellipodia. Nuclear accumulation of galectin-3, previously observed in proliferating 3T3 cells (Moutsatsos *et al.*, 1987), was not observed in proliferating T84 cells

<u>D.</u> <u>Rat recombinant galectin-4 enhances adhesion of T84 cells</u>.

Specific accumulation of galectin-4 in the cell-substrate contact sites in attaching cells, in lamellipodia of growing cells, and at the basal membrane of confluent monolayers suggests its involvement in cell-substrate adhesion. As the first attempt to investigate this possibility, we tested the ability of surface-adsorbed rat recombinant galectin-4 to support adhesion of T84 cells, using the procedure of McClay *et al.* (1981).

Adhesion of T84 cells to coated wells was examined by procedures described in detail by McClay *et al.* (1981) and Lotz *et al.*, (1989). Every other 8-well row in the flexible microtiter assay plate (MicroTest III, Becton Dickinson, Oxnard, CA) was treated with solutions of laminin (5 μg/cm2), BSA (2%) or rat recombinant galectin-4 (5 μg/cm2) in PBS. After addition of 50 μl to the wells, the plates were covered with Saran Wrap and incubated for 12 hours at room temperature. Coated

wells were washed 3 x 10 min. with PBS, and nonspecific adhesion sites were blocked by incubation with 2% BSA in PBS for 2 hours at room temperature, followed by 3 x 10 min. washes with PBS.

Subconfluent T84 cell cultures were labeled with 10 μCi/ml of [35S]-Met/Cys as above, rinsed quickly with Ca^{2+}- and Mg^{2+}-free PBS, and gently dissociated to single cells and small cell clusters by incubation for 1 min. at 37°C with 0.05 % trypsin solution containing 0.02% EDTA, prewarmed to 37°C. The cells were rinsed once in the growth medium, and resuspended in the fresh growth medium to 4 x 10^4 cells/ml. The viability of cells as assessed by trypan blue exclusion was > 99%. 300 μl of cell suspension was added to fill each well, and wells were sealed with an adhesive tape, avoiding formation of air bubbles. The plates were centrifuged at 24.9 x g for 8 min. at 4°C to bring the cells to contact with the substrate ("spin on"), and after incubation for 30 min. at 37°C, the plates were inverted and centrifuged at 100 x g, for 8 min. at 4°C, to dislodge the nonadhering cells from the substrate ("spin off"). The plates were then quick-frozen in a dry ice-ethanol mix, and the bottom 3 mm of each well were excised, and the adhering cells were quantified by scintillation counting in the presence of scintillation fluid. The total amount of applied cells (100%) was obtained by quantitation of cells remaining at the bottoms of BSA-coated wells immediately after the first centrifugation ("spin on").

For these studies, purified recombinant rat galectin-4 was prepared. E. coli producing full length rat galectin-4 were generated using the pET system (Novagen, Madison, WI) as described by Oda *et al.* (1993). The culture, extraction, and lectin-purification was done as described by Oda *et al.* (1993), except for induction at lower temperature and French press lysis to circumvent the insolubility of the galectin-4. In brief, 1 liter of Luria broth (LB) with 50 mg ampicillin was inoculated with a 15 ml overnight culture of the galectin-4 expressing E. coli, and then shaken for 3 hours at 37°C. The temperature was lowered to 30°C, IPTG (100 mg) added, and the incubation continued for another 3 hours. The bacteria were harvested by centrifugation and the pellet was either frozen or processed immediately. The bacterial pellet was resuspended in 15 ml of PBS containing 4 mM

b-mercaptoethanol and 2 mM EDTA (ME-PBS), and 1 mM PMSF on ice and lysed three times with a French press, the lysate was centrifuged at 100,000 x g for 45 min. The galectin was purified from the clear supernatant by affinity chromatography on lactosyl-Sepharose. The yield of active soluble galectin-4 was about 20 mg per liter of bacterial culture. The galectin was stored at 4°C in ME-PBS containing 150 mM lactose. Before use, an aliquot was chromatographed on a Superdex 75HR column (Pharmacia, Piscataway, NJ) in ME-PBS to remove lactose.

The subconfluent T84 monolayers were dissociated to single cells and small cell clusters (such as shown in Fig. 7a), and brought into contact with microtiter wells coated with either rat recombinant galectin-4, laminin or BSA. After 30 min. of incubation at 37°C, followed by centrifugation of the inverted microtiter plates at 100 x g for 8 min. at 4°C, about 19% of all cells remained attach to galectin-4, compared to less than 1 % of cells bound to BSA, and 65% cells attached to laminin (Fig. 9). Hence, galectin-4 supports significant T84 cell adhesion, implying that it interacts with one or more receptors at the T84 cell surface.

F. Discussion

These studies describe the location of two very abundant galectins, galectin-3 and -4, in human adenocarcinoma T84 cells in the subconfluent state and after differentiation into a polarized epithelium. At any stage of cell differentiation, the two galectins segregate from each other. In confluent polarized cells they accumulate at opposite poles, with galectin-3 at the apical, and galectin-4 at the basal membrane. In subconfluent cells, they concentrate in different parts of lamellipodia, with an accumulation of galectin-4 at the leading edge, and galectin-3 localized more proximally. Their segregation implies association of the two galectins with different intracellular ligands, and different functions. The specific accumulation of galectin-4 at sites of cell-substrate contact, and induction of T84 cell adhesion by recombinant galectin-4 suggest involvement of this protein in cell-substrate interactions.

Concentration of these galectins in defined subcellular areas has been observed earlier both in cultured cells and in tissues. Apical localization of galectin-3

in T84 cells agrees well with the similar finding in other polarizing epithelial cell lines (Lindstedt *et al.*, 1993, Sato *et al.*, 1993) and in kidney epithelium (Foddy *et al.*, 1990). Galectin-3 was also found in the nucleus of some cells such as normal colon epithelium (Lotz *et al.*, 1992) and 3T3 cells (Moutsatsos *et al.*, 1987), but not colon cancer cells (Lotz *et al.*, 1992) in agreement with the present observation for the T84 colon carcinoma cells. Galectin-4 has previously been found in globular structures at the cell periphery corresponding to areas of adherens junctions in oral epithelium (Chiu *et al.*, 1992), and in aggregates at the apical membrane of lumenal cells in esophageal epithelium (Wasano and Hirakawa, 1995).

The natural ligands for galectins are thought to be β-galactosides, and the differences in their binding specificity for complex saccharides (Leffler and Barondes, 1986; Leffler et al., 1989; Oda *et al.*, 1993) are sufficient to explain differential targeting. However, since β-galactosides are not detectable in the cytoplasm, the intracellular galectin ligands must be of another nature, interacting either with the carbohydrate-binding sites or with other, as yet unidentified, binding sites on the galectin surface. Indeed, interaction between galectin-3 and a nuclear glucose-binding protein of 70 kDa (CBP70) has been demonstrated (Seve et al., 1994), and most recently galectin-3 was shown to associate with Bcl-2, a well known intracellular suppressor of apoptosis (Yang *et al.*, in press). Both of these associations occur via protein-protein interactions, but are inhibited by lactose, indicating involvement of a carbohydrate-binding site. The association of galectin-3 with cytoplasmic and nuclear ribonucleoprotein complexes reported by Wang *et al.* (1992) is not inhibited by lactose, indicating involvement of another site. To date, there is no information about intracellular ligands for galectin-4.

Two observations made during immunocytochemistry experiments suggest that one component of cellular galectin complexes might be a lipid. First, detergent treatment of PLP-fixed cells removed the galectins from many locations, suggesting that they were associated with readily extractable components, such as lipids. Secondly, the galectins remained soluble and active after exposure to methanol, an unusual property among proteins, indicating a remarkable stability in interactions

with a hydrophobic environment.

The nature of the large apical saponin-resistant accumulations of galectins (Fig. 4d), and their functional relationship to cellular activity are unknown. These accumulations, which are less than 1 μm thick and between 2-10 μm in diameter are present at or beneath the apical surface of 40-80% of confluent T84 cells, either intra- or extracellularly. It is tempting to speculate that these galectin accumulations would create a specific microenvironment, and thus influence biophysical properties of the cytoplasmic membrane. One functional consequence of such accumulations could be structural, by creating conditions to stabilize and reinforce the epithelial cell membranes designed to function under constant mechanical stress. This observation agrees with the proposed role of galectin-4 in stabilizing the apical (lumenal) membranes in rat esophageal epithelium (Wasano and Hirakawa, 1995).

Our attempts to localize galectin-4 in adherens junctions of T84 cells were unsuccessful. Highly concentrated galectin-4 was visible at the lateral membrane in the form of large patches only after removal of Ca^{2+} with EGTA (Fig. 5), and only at the membranes of cells surrounding breaks in the monolayer, whereas in these conditions, double immunostaining showed a more even distribution of cytokeratin and the intracellular domain of E-cadherin in every cell. It remains to be determined if galectin-4 is translocated to lateral membrane sites as a result of Ca^{2+}-depletion, or if, instead, some proteins are removed in the process of monolayer dissociation, unmasking galectin-4 already present there. Even if galectin-4 is translocated to, or exposed at the lateral membrane only under certain conditions, it may still play a role in the physiological regulation of monolayer integrity.

The accumulation of galectin-4 at the basal membrane of confluent T84 cells, in the leading edge of lamellipodia in subconfluent T84 cells, and at cell-substrate contact sites in freshly seeded cells suggests a function in cell-substrate interactions. This view is further supported by the following: 1) the two tandem carbohydrate binding domains of galectin-4 (Fig. 2) should enable it to cross-link glycoconjugates at the cell surface and/or in the extracellular matrix; 2) evidence that galectin-4 is externalized at the basolateral surface of polarized cells including T84

(Huflejt, M.E., *et al.*, 1995, *J.Cell.Biochem.* 19B, p.20); 3) the observation that surface-adsorbed recombinant galectin-4 induces adhesion of T84 cells (Fig. 9). Furthermore, the sharply demarcated galectin-4 accumulations observed at the lateral cell surfaces at the sites of breaks in the monolayer after Ca^{2+} depletion, and at sites of cell-substrate adhesion invites the speculation that these very structures help reattach and spread cells in a disintegrated monolayer. Indeed, formation of the lamellipodia-like extensions has been observed *in vivo* during restitution of the intestinal epithelium (Rutten and Ito, 1983; Moore *et al.*, 1989), and after wounding of a T84 monolayer (Nusrat *et al.*, 1992). Therefore, galectin-4 is likely to play a role in the maintenance of epithelial integrity, and in the epithelial wound healing process.

Example 2

Detection of Human Galectin-4 in Tumors

Sixteen primary breast carcinomas that had been fixed in 10% buffered formalin (24-36 hours) and embedded in paraffin were obtained. The cases containing normal, benign and invasive components were selected from the tissue files of the University of North Carolina Breast Cancer Study and from Scripps Clinic and Research Foundation. All of the human breast tumor specimens analyzed contained both *in situ* and invasive component.

Tissue was processed for the immunohistochemical localization of galectin-4 using immunohistochemical localization. Rabbit anti-galectin-4 serum was raised against the C-terminal domain of rat intestinal galectin-4 as described (Oda *et al.*, 1993). To localize galectin-4 in human tissues, freshly cut or stored at 4°C 5 µm thick paraffin sections were used. The sections were deparaffinized in xylene (Fisher, Fair Lawn, NJ) through ethanol to PBS pH 7.2 followed by double distilled H_2O, then heat-treated in a microwave oven in 10 mM sodium citrate/NaOH pH 6.0 buffer at high power setting for 5 min and low power setting for 10 min (antigen retrieval method, Shi *et al.*, 1991). Sections were then cooled in the same buffer at room temperature, rinsed with double distilled H_2O, and after quenching of

endogenous peroxidase activity by incubation in 3% H_2O_2 in Ca^{2+}, Mg^{2+}-free PBS containing 0.1% Tween-20 and 0.1% NaN_3 for 20 min, nonspecific antibody binding was blocked by incubation in 10% normal goat serum (Vector Laboratories , Burlingame, CA) for 1 hour at room temp. Sections were then incubated with primary rabbit anti-galectin-4 antibody (1:2000 dilution) for 1 hour at room temp., rinsed in PBS containing 0.1% Tween-20, incubated with HRP-conjugated goat anti-rabbit antibody (BioRad Laboratories , Hercules, CA) (1:300 dilution) for 30 min at room temp., and the color reaction was developed for 5 min with the Liquid DAB Substrate (BioGenex, San Ramon, CA). After counterstaining with hematoxylin, sections were dehydrated in a series of ethanol solutions followed by xylenes and then mounted in Permount (Fisher, Dallas, TX).

While the tumor cells of both *in situ* and invasive tumor tissues have shown strong positive immunostaining with anti-galectin-4 antibody, the adjacent normal tissues remained unstained (Fig. 10). Galectin-4 was also absent in lymphocyte aggregates, endothelium, and in stroma of benign tumors. However, galectin-4 was present in both stroma and in endothelium of highly invasive tumors showing large quantities of disseminated epithelial tumor cells.

Reactivity of the breast tumor cells with the anti-galectin-4 antibodies was very specific, and it was possible to detect small nests of tumor cells or even single tumor cells in unexpected locations such as adipose tissue and aggregates of small infiltrating lymphocytes (Fig. 10G and Fig. 10H). Galectin-4 was also detected in certain cells within epithelium of cancerized lobules. In each such case, cells expressing galectin-4 had enlarged nuclei and were slightly bigger than neighboring, galectin-4-negative cells.

In all cases of human breast tissues examined in this study (n=16), galectin-4 was present in tumor cells, and absent in adjacent morphologically normal tissues. Such specific expression pattern together with apparent abundance of galectin-4 in tumor cells suggests that this protein may be used as a marker for breast tumor, with several potential diagnostic applications. The abundance of galectin-4 and high sensitivity of the assay offers the possibility of marked increase in sensitivity of

detection of tumor cells in: **1.** Breast tissue biopsies, including fine needle biopsies in which only a very small amount of tissue is sampled. **2.** Occult metastases, which are easy to miss during routine examination; this is particularly important in case of early tumor dissemination. **3.** Lymph nodes.

The specificity of galectin-4 expression indicates the possibility of using this protein in differential diagnosis of metastatic adenocarcinomas in tissues which normally do not express galectin-4, such as lung.

At the present time there is a dilemma in the field of diagnostic surgical pathology. The current view is that at least some breast carcinomas arise out of the progression of an intraductal proliferative process that may begin with hyperplasia and eventuate in an *in situ* carcinoma. The clinical problem relates to our lack of understanding of this process and our inability to precisely determine the point at which the intraductal proliferative process should be regarded as intraductal carcinoma rather than atypical intraductal hyperplasia. Expression of galectin-4 at early stages of proliferative breast disease in tumor tissues suggests that the expression pattern of this protein may help to distinguish between a benign process and a malignant process, potentially indicating patients with the higher risk of progression to malignancy. The expression of galectin-4 in individual cells of otherwise morphologically normal ducts in the tumor vicinity may help to identify cancerized lobules, and thus may be used to assess the extent of malignancy more accurately. In addition, the presence of galectin-4 in endothelium and in the lumen of blood vessels in tumor tissues suggests the presence of this protein in the blood of breast cancer patients: in this case, the assay monitoring blood levels of galectin-4 can be developed as a diagnostic screening tool. The expression of galectin-4 in tumor but not in normal tissues indicates possibility for the development of novel therapeutic strategies.

Example 3

Transgenic Mice Expressing Human Galelctin-4

Transgenic mice that can produce human galectin-4 are made according to methods well known in the art. For example, young female mice are injected with hormones to induce superovulation, and are mated. The one-celled embryos are collected, and the pronuclei injected with a purified DNA solution containing the constructs which encode human galectin-4. The injected eggs are cultured briefly, and re-implanted into pseudopregnant female mice. Genomic DNA is prepared from each progeny, and analyzed by PCR or Southern blot to determine the mouse's genotype. Mice who carry the transgene are subsequently mated in order to produce a line of mice.

Using these method mice are produced which contain transgenes encoding human galectin-4. In order to target the expression to specific cell types, transgenic mice are produced which carry genes encoding human galectin-4 under the control of specific promoters. Examples of such constructs are genes encoding SEQ ID NO:2 under the control of the rat whey acidic protein regulatory sequences (Yarus, S., et al., 1997, "The carboxy-terminal domain of human surfactant protein B is not required for secretion in milk of transgenic mice," *Front. Biosci.* 2:A1-A8) , or the mouse mammary tumor virus (MMTV) LTR (Mediavilla, M.D., et al., 1997, "Effects of melatonin on mammary gland lesions in transgenic mice overexpressing N-ras proto-oncogenes," *J. Pineal Res.* 22:86-94), which specifically target the expression of human galectin-4 to breast tissue.

Although the invention has been described with reference to the presently preferred embodiment, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

What is claimed is:

1. An isolated polynucleotide encoding human galectin-4.

2. The polynucleotide of claim 1, having the amino acid sequence as set forth in SEQ ID NO:2.

3. The polynucleotide of claim 2, having the sequence as set forth in SEQ ID NO:1

4. An isolated polynucleotide selected from the group consisting of:
   a) SEQ ID NO:1;
   b) SEQ ID NO:1, wherein T can also be U;
   c) nucleic acid sequences complementary to SEQ ID NO:1;
   d) fragments of a), b), or c) that are at least 15 bases in length and that will hybridize to DNA which encodes human galectin-4 as set fort in SEQ ID NO:2.

5. An isolated polynucleotide having 50% or greater sequence identity with SEQ ID NO: 1.

6. The polynucleotide of claim 1, further comprising an expression control sequence operatively linked to the polynucleotide encoding human galectin-4.

7. The polynucleotide of claim 3, wherein the expression control sequence is a promoter.

8. The polynucleotide of claim 7, wherein the promoter is tissue specific.

9. An expression vector containing the polynucleotide of claim 1.

10. The vector of claim 9, wherein the vector is a viral vector.

11. A host cell containing the vector of claim 9.

12. The host cell of claim 11, wherein said host cell is a eukaryotic cell.

13. The host cell of claim 11, wherein said host cell is a prokaryotic cell.

14. Substantially purified human galectin-4 polypeptide.

15. The polypeptide of claim 14, having the amino acid sequence as set forth in SEQ ID NO:2.

16. An isolated polypeptide having an amino acid sequence which is at least 85% identical to SEQ ID NO:2.

17. An isolated polypeptide which is a fragment or analog of SEQ ID NO:2.

18. An antibody which binds human galectin-4.

19. The antibody of claim 18, wherein the antibody is monoclonal.

20. The antibody of of claim 19, wherein the monoclonal antibody is a humanized monoclonal antibody.

21. A method for detecting a human galectin-4-associated disorder in a subject, comprising:

    contacting a sample from the subject suspected of having a human galectin-4-associated disorder with a reagent that binds to human galectin-4,

    and detecting binding of the reagent to human galectin-4.

22. The method of claim 21, wherein the sample is nucleic acid.

23. The method of claim 22, further comprising amplifying the nucleic acid of the sample prior to contacting a sample from the subject suspected of having a human galectin-4 disorder with a reagent that binds to human galectin-4.

24. The method of claim 21, wherein the sample is breast tissue, blood, plasma, serum, or urine.

25. The method of claim 21, wherein the disorder is cancer.

26. The method of claim 21, wherein the disorder is breast cancer.

27. The method of claim 21, wherein the reagent is an antibody which binds to human galectin-4.

28. The method of claim 21 wherein the reagent is a polynucleotide which encodes human galectin-4 or fragments thereof.

29. The method of claim 21, wherein the detecting is *in vivo*.

30. The method of claim 21, wherein the detecting is *in vitro*.

31. The method of claim 27 or 28, wherein the reagent is detectably labeled.

32. The method of claim 31, wherein the detectable label is selected from the group consisting of a radioisotope, a fluorescent compound, a bioluminescent compound and a chemiluminescent compound.

33. A method of treating a subject having or at risk of having a human galectin-4-associated disorder comprising administering to the subject a therapeutically effective amount of a reagent that suppresses human galectin-4.

34. The method of claim 33, wherein the disorder is cancer.

35. The method of claim 33, wherein the disorder is breast cancer.

36. The method of claim 33, wherein the agent is an antisense oligonucleotide that hybridizes to a human galectin-4 nucleic acid.

37. The method of claim 33, wherein the agent is an anti-human galectin-4 antibody.

38. The method of claim 37, wherein said anti-human galectin-4 antibody is a humanized antibody.

39. The method of claim 37, wherein the antibody is administered within a dose range between about 0.1/kg to about 100 mg/kg.

40. The method of claim 37, wherein the antibody is formulated in a pharmaceutically acceptable carrier.

41. A formulation for administration of an antibody characterized by its ability to bind human galectin-4, comprising a therapeutically effective amount of an

antibodycharacterized by its ability to bind human galectin-4, and a pharmaceutically acceptable carrier.

42. The formulation of claim 41, wherein the carrier is a liposome.

43. A transgenic nonhuman animal having a phenotype characterized by expression of human galectin-4, otherwise not naturally occurring in the animal, the phenotype being conferred by a transgene contained in the somatic and germ cells of the animal, the transgene comprising a nucleic acid sequence which encodes human galectin-4 polypeptide.

44. The transgenic nonhuman animal of claim 43, wherein the animal is a mouse.

45. A method for determining the prognosis of a subject diagnosed with cancer, comprising:

obtaining a sample from said subject,

determining the presence or absence of the expression of human galectin-4 in said sample from said subject, and

correlating the presence or absence of human galectin-4 with the prognosis of said subject.

46. The method of claim 45, wherein said cancer is adenocarcinoma.

47. The method of claim 45, wherein said cancer is breast cancer.

48. The method of claim 45, wherein said cells are breast cells.

49. A method of determining susceptibility of a subject with a tumor with to metastases, comprising:

obtaining a sample from said subject,

determining the presence of absence of human galectin-4 in said sample from said subject, and

correlating the presence or absence of human galectin-4 in said subject with the ability of the tumor to metastasize.

50. The method of claim 49, wherein said tumor is an adenocarcinoma.

51. The method of claim 49, wherein said tumor is a tumor of the breast.

52. The method of claim 49, wherein said cells are contained in a biopsy sample from said subject.

53. A method of determining the presence of metastases in a subject, comprising:

obtaining a sample from said subject,

determining the presence or absence of the expression of human galectin-4 in the cells of said subject,

correlating the presence or absence of the expression of human galectin-4 in the cells of said subject with the ability of said cells to metastisize.

54. A kit useful for detecting the presence of human galectin-4 in a sample from a subject having a human galectin-4-associated disorder, the kit comprising: carrier means being compartmentalized to receive in close confinement therein one or more containers comprising a container containing an antibody which specifically binds to human galectin-4.

55. A kit useful for the detection of a target nucleic acid sequence in a sample from a subject having a human galectin-4-associated disorder, wherein the presence of the target nucleic acid sequence in the sample is indicative of having or predisposed to having a human galectin-4-associated disorder, the kit comprising:

carrier means being compartmentalized to receive in close confinement therein one or more containers comprising a container containing oligonucleotides which hybridize to human galectin-4 nucleic acid sequences.

56. A method of identifying a compound which stimulates or inhibits human galectin-4 expression, comprising:

incubating the compound and a cell of interest under conditions sufficient to allow the compound to interact with the cell,

detecting the presence or absence of human galectin-4 said cell with a reagent, and

comparing the expression of human galectin-4 in said cell incubated with the compound with the expression of human galectin-4 in a control cell not incubated with the compound.

57. The method of claim 56, wherein the reagent is an anti-human galectin-4 antibody.

58. The method of claim 56, wherein the reagent is a polynucleotide which encodes which human galectin-4 polypeptide or fragments thereof.

59. The method of claim 56, wherein the reagent is detectably labeled.

60. The method of claim 59, wherein the detectable label is selected from the group consisting of a radioisotope, a fluorescent compound, a bioluminescent compound, and a chemiluminescent compound.

FIGURE 1

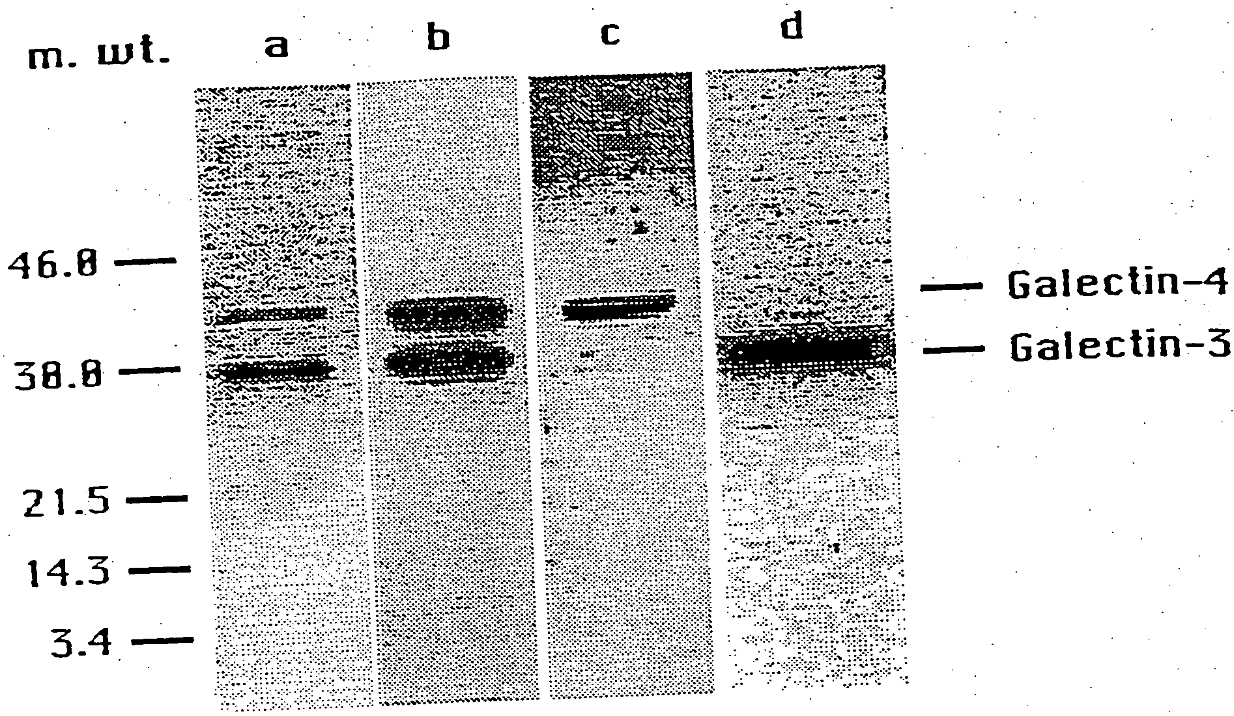


FIGURE 2A

```
127  CCGAAGCGGTTCTTCGTGAACTTTGTGGTTGGGCAGGATCCGGGCTCAGACGTCGCCTTC  186
 43  P  K  R  F  F  V  N  F  V  V  G  Q  D  P  G  S  D  V  A  F     62

187  CACTTCAATCCGCGGTTTGACGGCTGGGACAAGGTGGTCTTCAACACGTTGCAGGGCGGG  246
 63  H  F  N  P  R  F  D  G  W  D  K  V  V  F  N  T  L  Q  G  G     82

247  AAGTGGGGCAGCGAGGAGAGGAAGAGGAGCATGCCCTTCAAAAAGGGTGCCGCCTTTGAG  306
 83  K  W  G  S  E  E  R  K  R  S  M  P  F  K  K  G  A  A  F  E    102

307  CTGGTCTTCATAGTCATGGCTGAGCACTACAAGGTGGTGGTAAATGGAAATCCCTTCTAT  366
103  L  V  F  I  V  M  A  E  H  Y  K  V  V  V  N  G  N  P  F  Y    122

367  GAGTACGGGCACCGGCTTCCCCTACAGATGGTCACCCACCTGCAAGTGGATGGGGATCTG  426
123  E  Y  G  H  R  L  P  L  Q  M  V  T  H  L  Q  V  D  G  D  L    142

427  CAACTTCAATCAATCAACTTCATCGGAGGCCAGCCCCTCCGGCCCCAGGGACCCCCGATG  486
143  Q  L  Q  S  I  N  F  I  G  G  Q  P  L  R  P  Q  G  P  P  M    162

487  ATGCCACCTTACCCTGGTCCCGGACATTGCCATCAACAGCTGAACAGCCTGCCCACCATG  546
163  M  P  P  Y  P  G  P  G  H  C  H  Q  Q  L  N  S  L  P  T  M    182

547  GAAGGACCCCCAACCTTCAACCCGCCTGTGCCATATTTCGGGAGGCTGCAAGGAGGGCTC  606
183  E  G  P  P  T  F  N  P  P  V  P  Y  F  G  R  L  Q  G  G  L    202

607  ACAGCTCGAAGAACCATCATCATCAAGGGCTATGTGCCTCCCACAGGCAAGAGCTTTGCT  666
203  T  A  R  R  T  I  I  I  K  G  Y  V  P  P  T  G  K  S  F  A    222

667  ATCAACTTCAAGGTGGGCTCCTCAGGGGACATAGCTCTGCACATTAACCCGCATGGCAAC  726
223  I  N  F  K  V  G  S  S  G  D  I  A  L  H  I  N  P  H  G  N    242

727  GGTACCGTGGTCCGGAACAGCCTTCTGAATGGCTCGTGGGGATCCGAGGAGAAGAAGACT  786
243  G  T  V  V  R  N  S  L  L  N  G  S  W  G  S  E  E  K  K  T    262

787  ACCCACAACCCATTTGGTCCCGGACAGTTCTTTGATCTGTCCATTCGCTGTGGCTTGGAT  846
263  T  H  N  P  F  G  P  G  Q  F  F  D  L  S  I  R  C  G  L  D    282

847  CGCTTCAAGGTTTACGCCAATGGCCAGCACCTCTTTGACTTTGCCCATCCGTCTCGGGCC  906
283  R  F  K  V  Y  A  N  G  Q  H  L  F  D  F  A  H  P  S  R  A    302

907  TTCCAGAGGGTGGACACATTGGAAATCCAGGGTGATGTCACCTTGTCCTATGTCCAGATC  966
303  F  Q  R  V  D  T  L  E  I  Q  G  D  V  T  L  S  Y  V  Q  I    322

967  TAATCTATTCCTGGCATAACTCATGGGAAAACAGAATTATCCCCTAGACTCCTTTCTAAG  1026

1027 CCCCTAATAAAATGTCTGAGGGTGTC  1052
```

FIGURE 2B

```
hGal-4                                                              PKRFFVHF
rGal-4   MAYVPAPGYQ PTYNPTLPYK RPIPGGLSVG MSIYIQGIAK DNMR..H...
pGal-4   MAFVPAPGYQ PTYNPTLPYY KPIPGGLRVG MSVYIQGVAH EHM.......

hGal-4   VVGQDPGSDV AFHFNPRFDG WDKVVFNTLQ GGKWGSEERK RSMPFKKGAA
rGal-4   A....E.A.I .......... ........M. S.Q..K..K. K....Q..HH      Domain-I
pGal-4   ....G..A.. .......... .......SQ. D....N..K. .....R.AP.

hGal-4   FELVFIVMAE HYKVVVNGNP FYEYGHRLPL QMVTHLQVDG DLQLQSINFI
rGal-4   .....M..S. ........T. .......... .......... ..E......L
pGal-4   ....IM.LP. ........D. ...F...I.V .L........ ..T.......

hGal-4   GGQPLRPQGP P-MMPP-YPG PGHCHQQLNS
rGal-4   ....AAS.Y. GT.TI.A..S A.YNPP.M..                              Link
pGal-4   ....APS--. GT.PN.G... ..KHN..PCN

hGal-4   LPTMEGPPTF NPPVPYFGRL QGGLTARRTI IIKGYVPPTG KSFAINFKVG
rGal-4   ..V.A...I. ......V.T. .......... ........L..A .NLI......
pGal-4   ..C...A... ......KT.. ....V..... V......S. ..LV......

hGal-4   SSGDIALHIN PRMGNGTVVR NSLLNGSWGS EEKKITHNPF GPGQFFDLSI
rGal-4   .T....F.M. ..I.DC-... ..YM...... ..R..PY... .A........      Domain-II
pGal-4   ....V..... ..LTE.I... ..Y...K..A ..R.SSF... A...Y.....

hGal-4   RCGLDRFKVY ANGQHLFDFA HPSRAFQRVD TLEIQGDVTL SYVQI
rGal-4   ...T.....F .........S .RFQA..... M...K..I.. .....
pGal-4   .......... .........S .RLSN..G.. .......... .....
```

Fig. 3

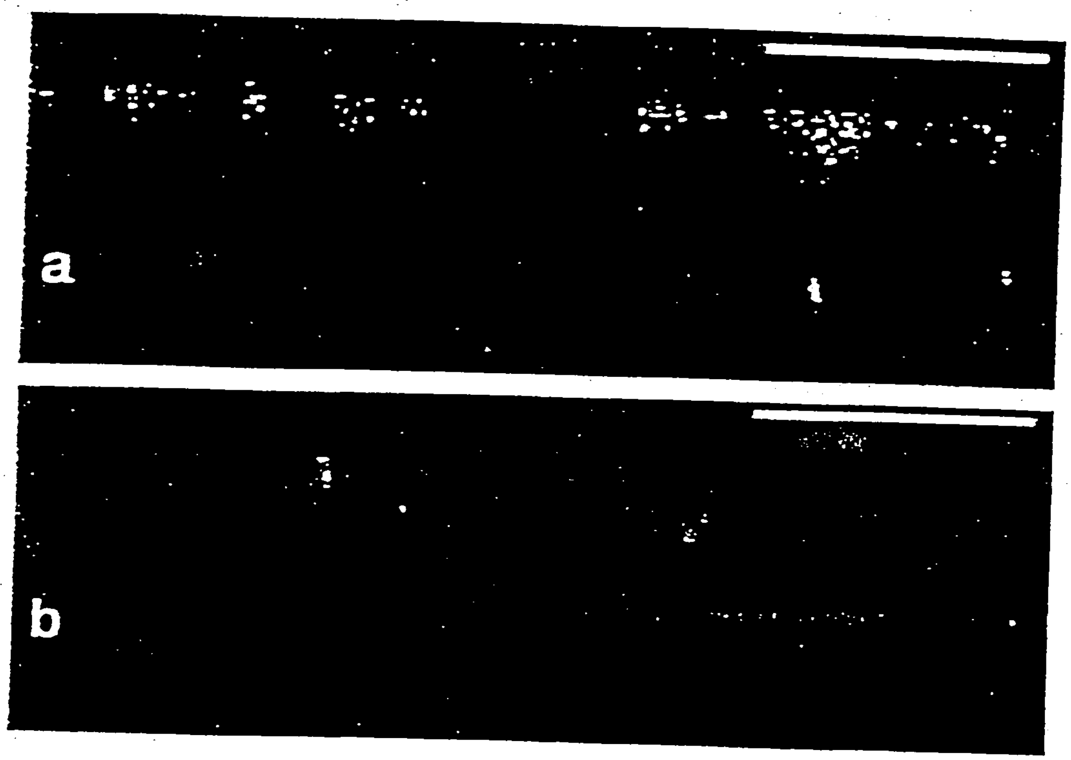


Fig. 4

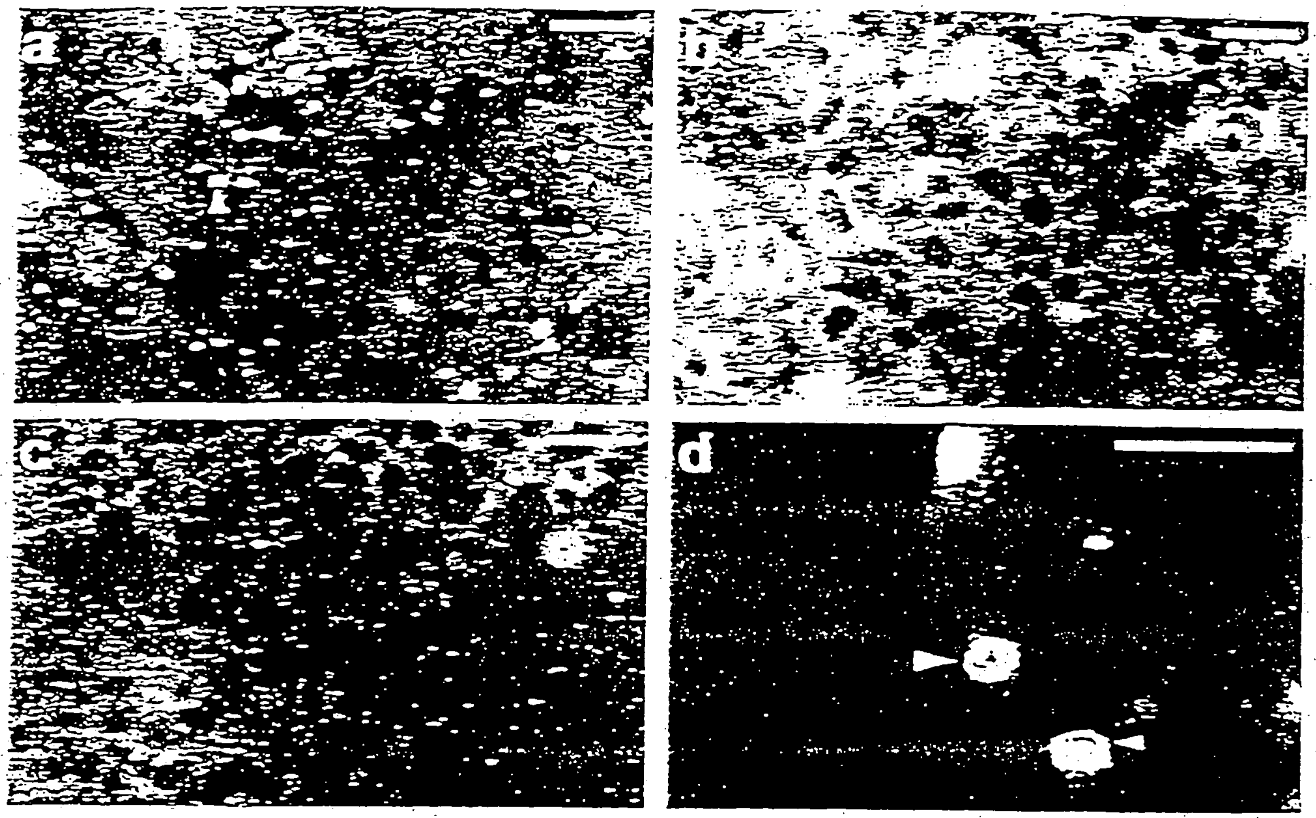


Fig. 5

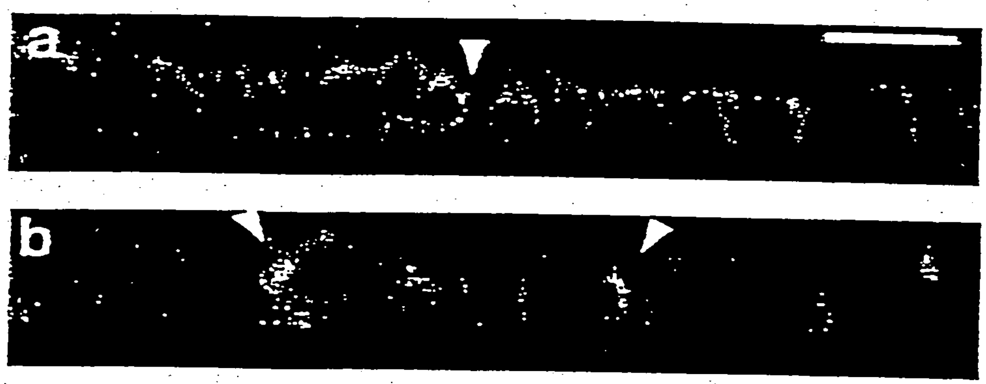


Fig. 6

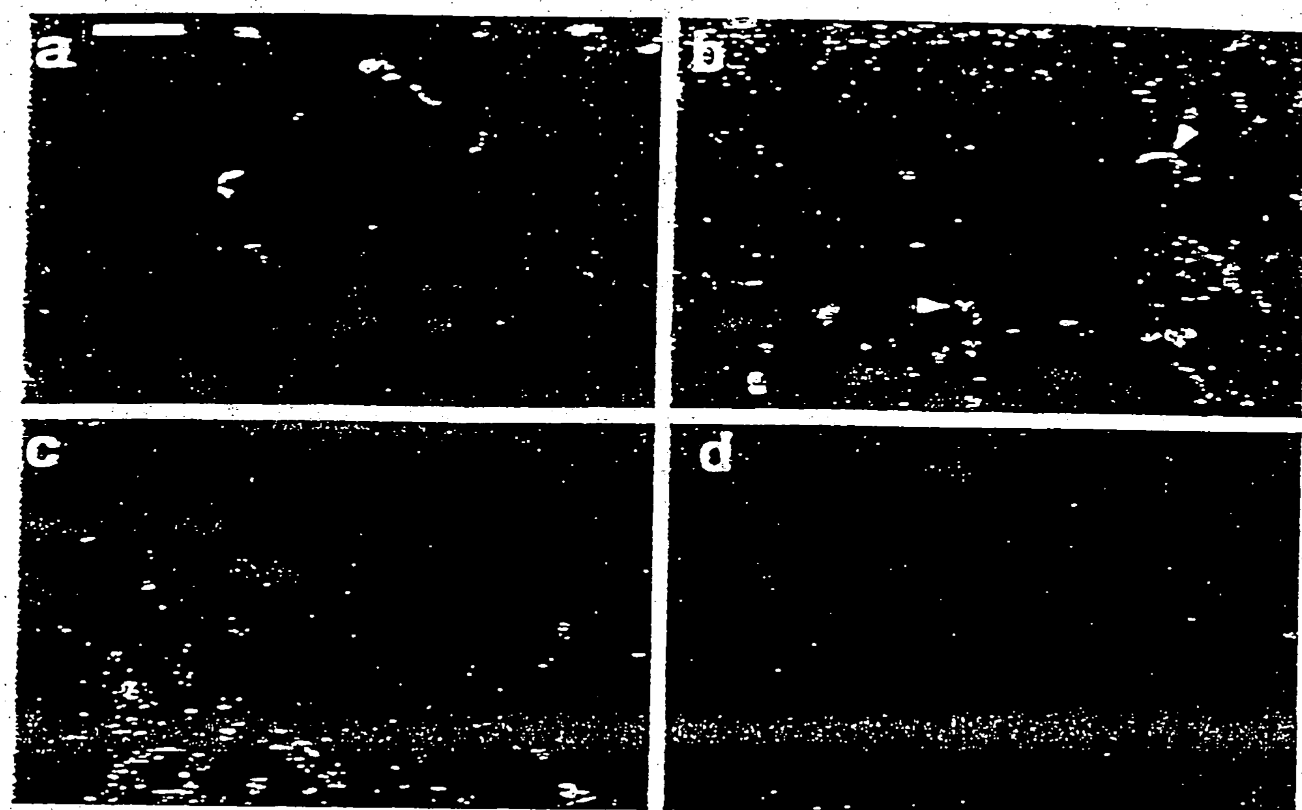


Fig. 7

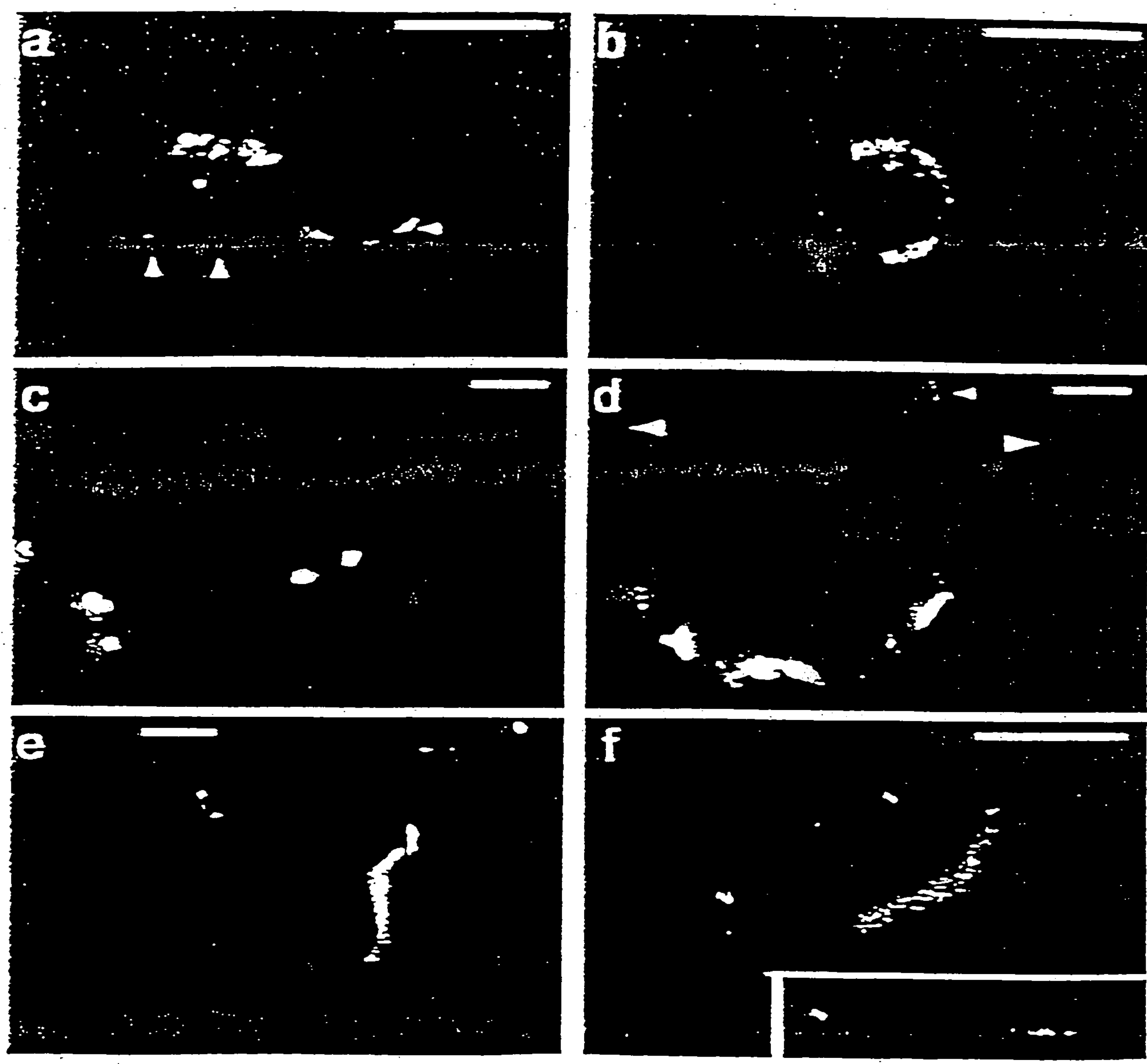


FIGURE 8

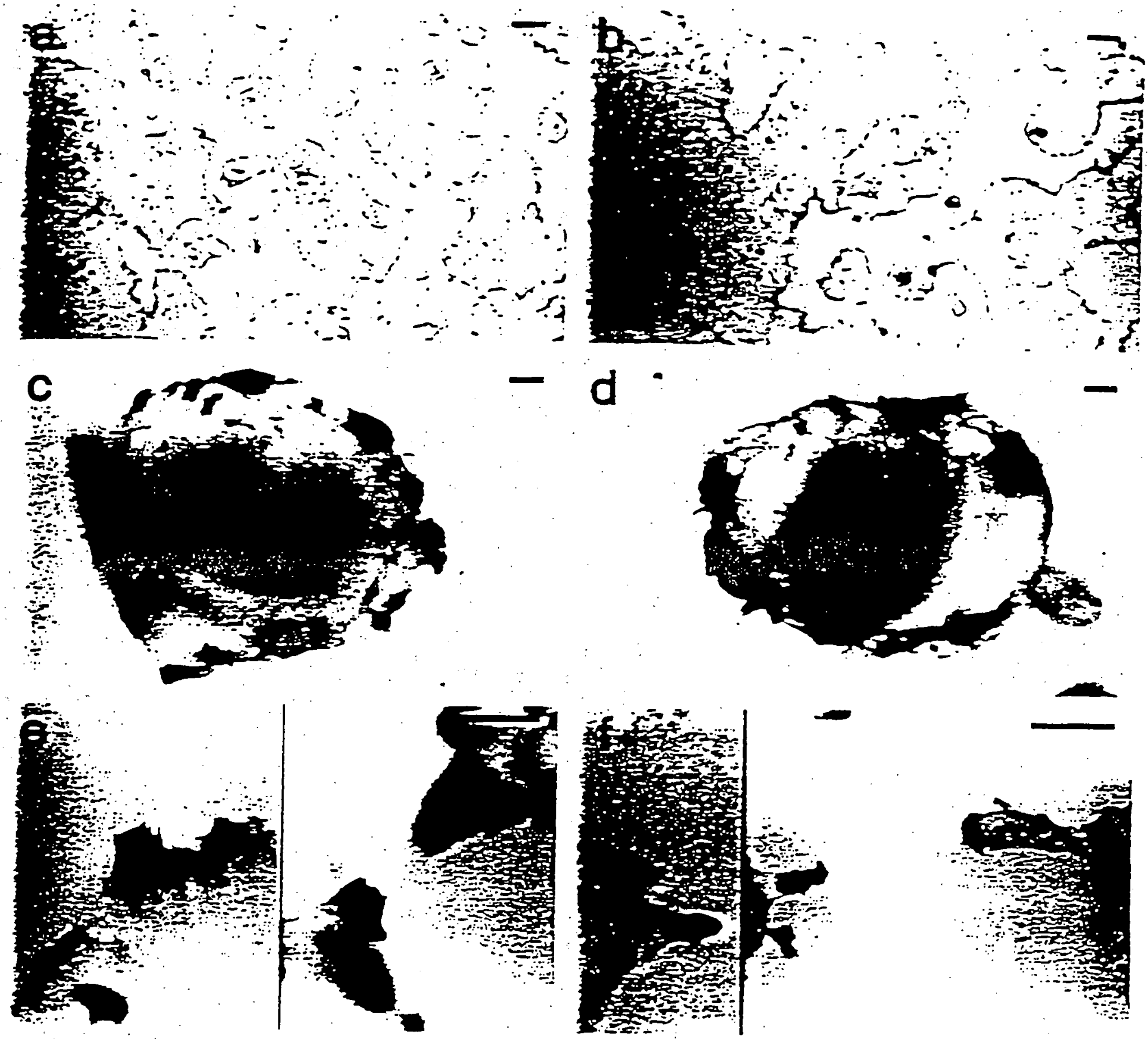


Fig. 9

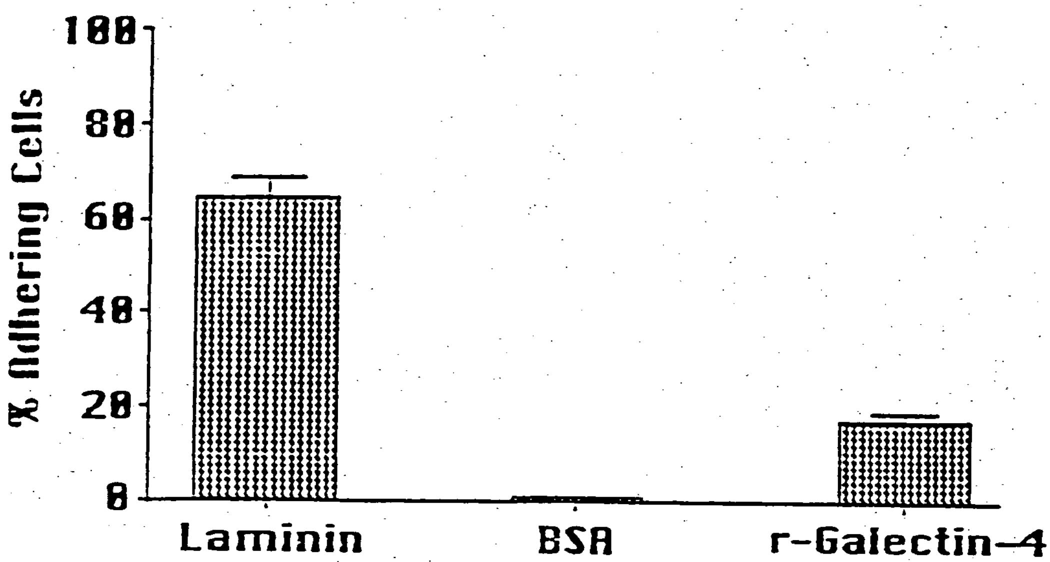
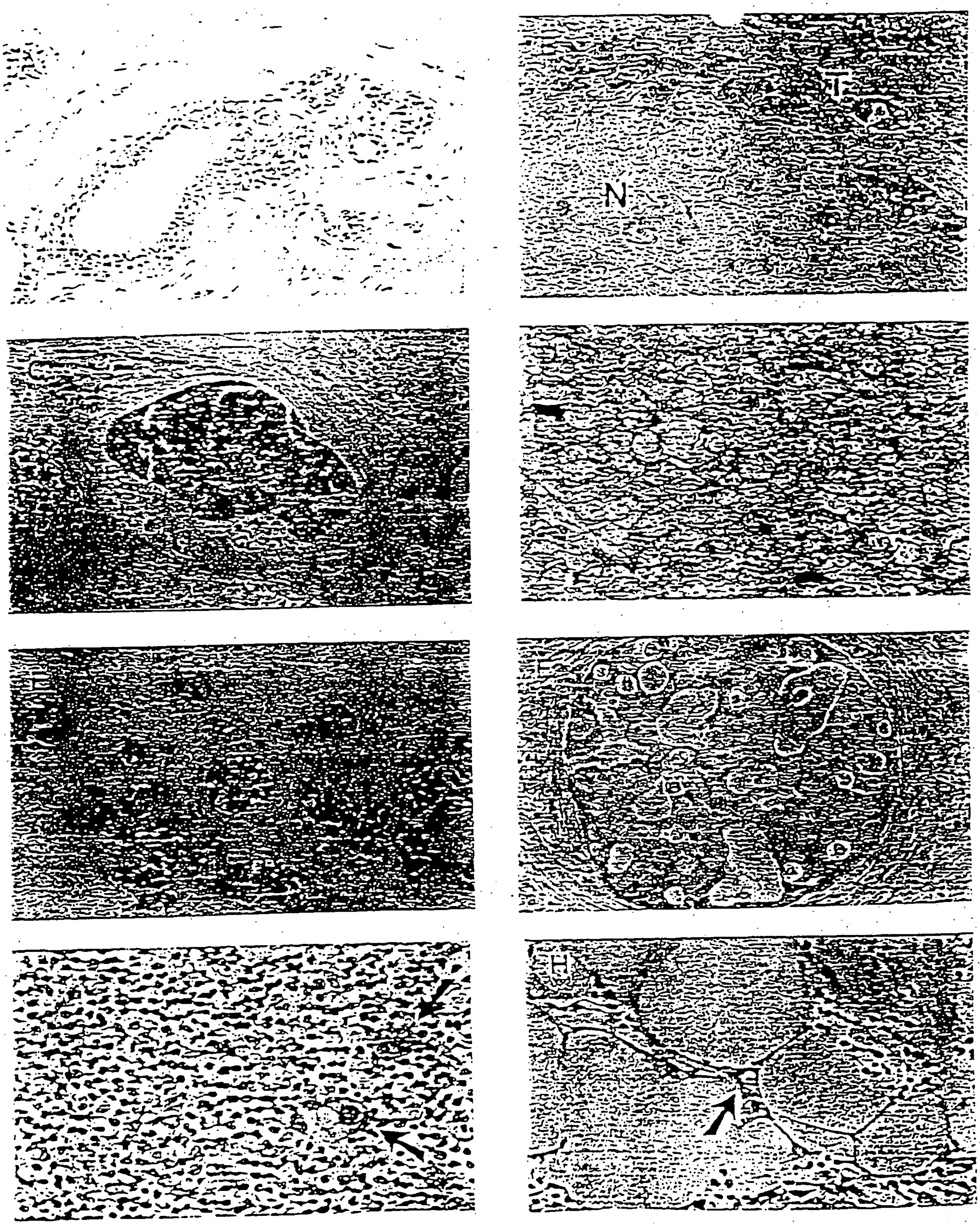


FIGURE 10



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/21807

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 530/350, 387.7, 387.9, 388.15, 389.7; 435/6, 7.23, 346; 424/9, 133.1, 138.1; 514/44; 800/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; DIALOG: file biochem
search terms: galectin, galectin-4, human, cancer, carcinoma, metasta?, neoplas?, epsilon-BP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	TARDY et al. Purification and Characterization of the N-Terminal Domain of Galectin-4 from Rat Small Intestine. FEBS Letters. 1995, Vol. 359, pages 169-172, especially the last full paragraph on page 169, and the paragraph bridging pages 171 and 172.	18 ---- 1-17, 19-60
X,P ---- Y,P	HUFLEJT et al. Strikingly Different Localization of Galectin-3 and Galectin-4 in Human Colon Adenocarcinoma T84 Cells. The Journal of Biological Chemistry. 30 May 1997, Vol. 272, No. 22, pages 14294-14303, especially page 14295 and Figure 2.	1-13 ------- 14-60

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 MARCH 1998	03 APR 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer TONI R. SCHEINER Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ---- Y,P	RECHRECHE et al. Cloning and Expression of the mRNA of Human Galectin-4, an S-Type Lectin Down-Regulated in Colorectal Cancer. European Journal of Biochemistry. 15 August 1997, Vol. 248, pages 225-230, especially Figure 2 and page 229.	1-13 ------ 14-60
A	KASAI et al. Galectins: A Family of Animal Lectins That Decipher Glycocodes. Journal of Biochemistry. January 1996, Vol. 119, pages 1-8.	1-60
A	MAI et al. A Novel β-Galactoside-Binding Lectin in Cultured Murine Lymphocytic Leukemia Cells. Journal of Biochemistry. September 1996, Vol. 120, pages 478-480.	1-60

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6): A61K 39/395; C07H 21/02, 21/04; C07K 14/435, 16/30; C12Q 1/68; G01N 33/574

A. CLASSIFICATION OF SUBJECT MATTER:
US CL : 536/23.1; 530/350, 387.7, 387.9, 388.15, 389.7; 435/6, 7.23, 346; 424/9, 133.1, 138.1; 514/44; 800/2